THE ROLE OF PREDICTION IN LANGUAGE LEARNING: EVALUATING ERROR-BASED THEORIES OF LANGUAGE ACQUISITION

Thesis submitted in accordance with the requirements

of the University of Liverpool for the degree of

Doctor in Philosophy by

Judit Fazekas

June 2020

DECLARATION IN HIGHER DEGREE THESES

University of Liverpool
Institute of Life and Human Sciences
Psychological Sciences

DECLARATION

This thesis is the result of my own work. The material contained in the thesis has not been presented, nor is currently being presented, either wholly or in part, for any other degree or qualification.

Signed ………………………………… (Candidate)

Date 31/01/2020 (Candidate)

ACKNOWLEDGEMENTS

First, I would like to thank my family, who have always encouraged me to keep going even when it meant getting further and further away from them. Anyu, Apu, Gyuri, Zsuzsi és Örs: Köszönök mindent! I would also like to thank the most recent addition to my family: Fraser. I'm grateful, that amongst other major sacrifices, you were willing to spend multiple lifetimes on British public transport aka our personal hell. Thank you for regularly going to the end of the world and back for my dreams!

I would also like to thank my extended office family for always being there for me and sharing the various pains and joys of getting through a PhD. Charlie, Dani, Eugenia, Jo, Kat, Katie, Lana, Moon and Hannah, this would never have been the same without you. I'm also extremely grateful to my non-office friends who listened to my endless discussion of academia with all the patience of the world.

None of this would have been possible without my supervisors Caro and Julian, who helped me find my own path but were always there in case I got lost! I'm also immensely grateful to my mentor at Davis, Tamara, who made me feel at home not only in her lab but in the entire continent. And my previous mentors, Csaba and Holly, for introducing me to psycholinguistics and for encouraging me not to give up on it.

Last but not least I would like to thank to the many children and adults, who have been tirelessly repeating sentences and describing weird pictures in exchange for some even weirder cards. I would also like to thank those helping with finding the participants and carrying out many, many bingo games: Beth, Claire and Ken, this would have not been possible without you!

This research was made possible by PhD Studentship awarded to Judit Fazekas by the Economic and Social Research Council.

RATIONALE FOR USING AN ALTERNATIVE THESIS FORMAT

This thesis has been prepared following the alternative paper format, in accordance with the guidelines provided by the University of Liverpool for including research papers in a doctoral thesis. This format was selected in order to facilitate the publication of this research in scientific journals. Chapters 2, 3 and 4 represent separate manuscripts and are structured in a way suitable for submission to peer-reviewed journals. The Introduction, Methods and Data analyses sections of Chapter 3.2 (Do children learn from their mistakes? Evaluating error-based theories of language acquisition) have already received in principle acceptance in a Registered Report format at the Royal Society Open Science journal.

The font and style of all chapters is consistent throughout the thesis. All citations are presented in a single bibliography at the end of the thesis and no reference section is provided after individual chapters. Otherwise, the chapters are presented in the same format as the manuscripts that would be submitted for publication. Thus, each chapter starts with a review of the relevant literature to introduce an informed reader to the topic and ends with a discussion of the implications of the results. As with a traditional thesis, a general introduction describing the background literature is provided in Chapter 1, and Chapter 5 offers an overall discussion and interpretation of the studies described across the three experimental chapters.

All chapters and the studies described in them were prepared in collaboration with my PhD supervisors, Professor Caroline Rowland and Professor Julian Pine. Chapter 4 (Processing surprising and predictable sentences – evidence from ERPs) has been carried out in the Cognitive Neuroscience of Language lab at the University of Davis, California and Professor Tamara Swaab, the head of the lab, has been acting as the project leader for the study described in this chapter and Kathryn Bousquet contributed to EEG training and study design. Furthermore, Andrew Jessop carried out the power analyses described in Chapter 3.

ABSTRACT

This project examines prediction's role in language acquisition by assessing the projections of error-based language acquisition theories, specifically the Dual-path model. It seeks to expand on results obtained by the prime surprisal paradigm and to develop novel studies building on this method.

In Chapter 2, we examined whether prime surprisal appears with passive as well as dative structures in adults. While we found a numerical, but non-significant prime surprisal effect in the dative study, no such effects appeared with passives. In Chapter 3, in a four-stage prime surprisal-based intervention study we examined whether surprising sentences lead to more learning than predictable ones. Here we found preliminary evidence for enhanced learning rates for abstract structures after surprising input with 5-6 year-old child participants. These long-term effects were stronger in younger children, although they were dissociated from immediate prime surprisal. We found no significant evidence for verb-dependent error-based learning effects in either group or for an abstract learning effect in the adult group alone. In Chapter 4 we compared the ERPs associated with predictable and surprising dative sentences in adult participants and found an enhanced N400 effect in surprising (as opposed to predictable) DOD (double object dative) sentences but we found no such effect in prepositional (PD) sentences. We also found no significant differences in the P600 region. While the N400 effect shows predictability-related processing differences, based on these results alone we cannot determine whether these differences are the result of active prediction or differences in the ease of integration.

Some of our key results (such as the replication of immediate prime surprisal effects in two of our studies, the preliminary evidence for enhanced learning rate with unexpected structures or the N400 effects) support error-based learning theories. However, other results (such as the dissociation of immediate prime surprisal and longer-term learning effects or the lack of P600 effect in surprising sentences) raise questions about the exact nature of the mechanism leading to these effects and warrant further examination.

CONTENTS

CHAPTER ONE: THESIS INTRODUCTION AND OUTLINE

1.1. Introduction

Linguistic ‘prediction’ is the phenomenon that listeners ‘predict’, or pre-activate, words before they appear in the speech stream. The potential implications of linguistic prediction have been receiving more and more interest in the language sciences in recent years. In particular, prediction’s role across different levels of linguistic processing has inspired various psycholinguistic theories regarding how humans learn and use language (e.g. Pickering & Garrod, 2013; Christiansen & Chater, 2016); This thesis will concentrate on one such theory in particular – an error-based theory of language acquisition – and it will evaluate its proposal that prediction is the driving force behind language acquisition.

Error-based theories of language acquisition propose that children (and adults) are constantly predicting upcoming words, and subsequently assessing the accuracy of these predictions (e.g. Chang, Dell & Bock, 2006; Ramscar, Dye & McCauley, 2013). When the prediction turns out to be incorrect, an error signal is generated, which is then used for learning. While these theories are particularly promising, especially as they can explain observational and experimental data that is challenging for alternative theories, their validity has yet to be directly proven, not least due to the limitations of the methods typically used to assess linguistic predictions.

This thesis targets prediction’s role in language acquisition by using a novel method: prime surprisal. The prime surprisal method builds on the well-established priming paradigm, and assesses the extent to which predictable and surprising utterances influence subsequent language production. The experiments conducted as part of this project combine prime surprisal with other experimental techniques to discover more about linguistic predictions and their role in language acquisition. The main advantage of this approach is that prime surprisal can directly target the main prediction of error-based learning theories: that un-predictable (surprising) input leads to a larger prediction error, and in turn more language change, than predictable input.

The first, introductory, chapter starts with a discussion of prediction's potential role in language acquisition and discusses the Dual-path model, the theory that operationalises prediction-led learning in syntactic acquisition. The chapter will continue with an overview of existing research on prediction, with an enhanced focus on developmental studies on the topic. It will discuss the strengths and limitations of the methods typically used in prediction research and consider the relevance of the results for error-based learning theories. The chapter will conclude with a discussion of prime surprisal, the main method used throughout the dissertation.

Chapter 2, the first experimental chapter, discusses two prime surprisal studies with adult participants. The first study is a replication of earlier prime surprisal studies involving the dative alternation, while the second study features the active-passive transitive alternation that has, to our knowledge, not yet have been examined in other published prime surprisal research. The aim of these studies is to uncover more about predictions and their role in changing subsequent language behaviour by assessing whether prime surprisal can be observed in sentences involving a wider variety of linguistic structures.

While Chapter 2 concentrates on short-term effects of prediction, the second experimental chapter, Chapter 3, features a study targeting the longer-term effects of being exposed to predictable versus surprising structures. This chapter describes a prime-surprisal-based dative study involving two participant groups of adults and 5-6 year old children respectively. This study assesses whether less predictable input leads to more long-term as well as more short-term language change. Discussion of the main study is preceded by the description of the pilot study featuring 5-6 year-old child participants that was used as a basis for the main study. A version of the Introduction, Methods and Analyses plan sections of the main study has received In Principle Acceptance in the Royal Society Open Science Journal as a Registered Report.

Chapter 4 contains the description of an event related potential (ERP) study featuring adult participants. The goal of this study was to gain more information about the on-line processing of predictable and surprising sentences. To achieve this, we converted the study described in Chapter 3 into an EEG study in order to replicate both the long- and short-term effects assessed by the earlier study and, crucially, to compare the neural correlates arising during the processing of predictable and expected sentences.

The final, discussion chapter summarises the findings discussed in Chapters 2 to 4 and considers their implications for error-based theories of language acquisition. It also discusses potential directions for future research.

1.2 Error-based theories of language acquisition and the Dual-path model

Prediction, the ability to pre-activate upcoming input before we encounter it, has been receiving widespread attention in various fields of linguistic cognition. Prediction is used as the principal element in several theories that seek to model different levels of linguistic processing. Some theories highlight the role of prediction in enabling swift turn-taking in dialogue (e.g. Levinson, 2016), while others focus on prediction's relationship with speech production (e.g. Pickering & Garrod, 2013). Most importantly for the purposes of this account, it has been proposed that prediction might also be the key driving force behind language acquisition (e.g. Chang et al., 2006; Ramscar et al., 2013).

Error-based theories of language acquisition have gained wide support for a variety of reasons. First, they propose an integrated model of language acquisition by highlighting the relationship between language acquisition and language processing. They propose that as each word is processed, it is compared to the word that was predicted by the language processing mechanism, and that any error-signal resulting from this comparison is the basis of language development. As processing is an integral feature of the learning mechanism, these theories assume that limitations in processing can influence acquisition. Secondly, and crucially for experimental studies, error-based theories describe a well-defined learning mechanism. While earlier theories of language acquisition focused on defining children's state of knowledge at different points of development (e.g. Hirsh-Pasek & Golinkoff, 1996; Olguin & Tomasello, 1993), error-based theories describe how children move from one knowledge state to the next. Another strength of these models is that they can explain developmental phenomena that were challenging for earlier language development theories. Error-based theories can model, for example, how children recover from overgeneralisation errors without receiving explicit corrections. For instance, an error-based noun-acquisition model proposed by Ramscar and colleagues (2013) suggests that when children predict the overgeneralised "mouses" form and end up hearing the correct "mice", the discrepancy between the prediction and the input leads to an error signal that weakens the associations between the plural of "mouse" and "mouses" but strengthens the associations with "mice". Eventually the associations between the plural of "mouse" and the adult-like "mice" form become stronger than the connections with "mouses", and children start producing – and predicting – the correct form.

Error-based models of language acquisition are typically connectionist models with a simple recurrent network architecture (Rumelhart, Hinton & Williams,

1986) that operates via back-propagation of error (Christiansen & Chater, 1999; Elman, 1990). In these models, forward spreading activation generates an activation pattern that determines the model's expectations. The difference between these expectations and the actual input (the error) is measured and is then used to change the weights in the network via back-propagation. As this mechanism gradually adjusts the models internal representations, the model's expectations continuously improve. Various error-driven theories concentrate on explaining the acquisition of different elements of language competence. For instance, Oppenheim, Dell and Schwartz (2010) model lexical retrieval in speech production using an error-based learning mechanism, while Ramscar and colleagues' aforementioned model concentrates on noun-acquisition. In this work, we focus on the acquisition of syntax, thus we will be discussing the Dual-path model, a connectionist frequency-based model that operationalises syntactic acquisition in predictive terms (Chang et al., 2006; Chang, Janciauskas & Fitz, 2012) in more detail.

The Dual-path model describes syntactic acquisition via error-based learning and operates on the basis that people are constantly trying to predict the next word in the speech stream. For instance, upon hearing "It was a windy day so the boy went out to fly a …" the listener would already be anticipating the word "kite" even before the word appears in the speech stream.

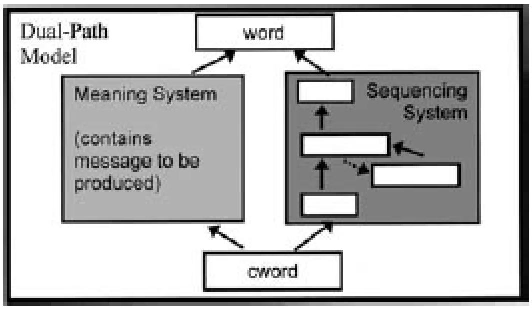

Figure 1.1. The two networks of the Dual-path model

(reproduced from Chang, Dell & Bock, 2006)

As the model's name suggests, the Dual-path model consists of two networks, a sequencing system (which is a simple recurrent network) and a hidden meaning network, see Figure 1.1. This dual architecture allows for the possibility that the same message might be expressed using different syntactic structures. For example, the sentences "Dora gave Boots a ball" and "Dora gave a ball to Boots" have a very similar meaning, but this meaning can be expressed using two different syntactic structures.

In this model, the syntax system leads to the generation of word by word predictions by taking into account the previous word (sequencing system) and the intended message (meaning system). Once the predicted word appears in the speech stream it is compared to the previously made prediction and any discrepancy between the expectation and what is heard leads to an error signal. This error signal is then used to make changes in the weights supporting syntactic knowledge in the system. As the syntax system goes through more and more modifications (based on the error signals that are generated), the closer it gets to the language it is processing and the more accurate the subsequent predictions become. In the case of language acquisition, children can get closer to adult-like syntactic knowledge with each linguistic exchange, eventually reaching an adult-like syntactic system. Importantly, the error-based learning mechanism stays ‘switched on’ during adulthood as well, and it can facilitate adaptation to different linguistic environments and speakers.

Error-based learning theories are especially promising for language acquisition research due to their integrated and efficient nature, high explanatory power and clearly defined architecture leading to testable predictions. Given the widespread interest in predictive accounts of language processing and acquisition, various studies have investigated prediction in child- and adulthood. However, several key questions remain unanswered, due in part to the challenges of assessing predictive language processing experimentally. The following sections will review existing research into prediction and the potential limitations of this literature.

1.3. Research into linguistic predictions

1.3.1. What do we mean by prediction?

Various fields of language sciences have shown considerable interest in the phenomenon of prediction, and the wide variety of approaches adopted by different groups of researchers has led to multiple parallel definitions of predictions. These definitions often disagree on key characteristics of this phenomenon, and the different assumptions that underpin these definitions can influence how studies are conceptualised. Multiple opinions exist with regards to the sources of information speakers use for their predictions, which linguistic levels are predicted (e.g. semantics and/or phonology), the number of consecutive words predicted, the number of parallel predictions speakers make, and whether these predictions are graded. Kuperberg and Jaeger (2016) provide a detailed discussion regarding these debates and the key differences in approach.

This account will also consider some of these distinctions in the following chapters, but in light of the divergences in usage it is important to first clarify what definition of ‘prediction’ is used throughout this work. Unless otherwise specified, this work considers predictions to be made one word at a time, consist of one prediction (and not several alternative predictions), and to be made on all linguistic levels. This definition is compatible with the definition used by the Dual-path model (Chang et al., 2006).

The subsequent sections will discuss the most important elements of the literature on linguistic prediction in more detail. A wide range of experimental paradigms have been used to assess predictions, typically by measuring the differences in processing predictable versus expected linguistic stimuli. These studies have shown that more predictable words are identified faster in both reading and listening comprehension than predictable words (e.g. Schwanenflugel & Shoben, 1985; Traxler & Foss, 2000). In this account, we will concentrate on discussing the literature most relevant to prediction’s role in language acquisition. As such, we will focus on results of experimental paradigms involving child participants (studies addressing turn taking and studies using the looking-while-listening paradigm), as well as certain methods that aim to addresses some of the shortcomings of the aforementioned studies (event-related potential studies).

1.3.2. Conversational turn-taking

Turn taking in adult dialogue is remarkably fast: adults often start speaking as little as 200 milliseconds after their partner finishes (De Ruiter, Mitterer, & Enfield, 2006), even though utterance planning typically takes much longer (for instance, picture naming can take over 600 ms; Indefrey & Levelt, 2004). A reason behind these short pauses might be that people predict when their conversational partner will stop speaking and prepare their responses in advance so they can start their utterance as soon as possible (e.g., Levinson, 2016). While there are alternative explanations for this phenomenon (e.g. Heldner & Edlund, 2010; Pickering & Gambi, 2018), the results of multiple studies support the idea that prediction is the driving force behind swift turn taking. For example, Magyari, Bastiaansen, De Ruiter, and Levinson (2014) showed that participants accurately predicted when recordings extracted from a corpus of Dutch telephone conversations would end, and their accuracy was even higher for more predictable endings. This result was also supported by event related potential (ERP) data showing an early (beginning at least 1250 ms before the end of the utterance) power decrease for predictable as opposed to unpredictable turns (see also De Ruiter et al., 2006 and Magyari & De Ruiter, 2012).

In contrast to the speedy turn taking of adults, children tend to be significantly slower and can take up to 600 ms to start their answer once their conversation partner stops speaking (Casillas, Bobb, & Clark, 2016; Stivers, Sidnell, & Bergen, 2018). This difference is even more remarkable when taking into account the fact that, in their pre-linguistic interactions, children tend to be much faster in turn-taking and they only start slowing down once they start using language (Hilbrink, Gattis, & Levinson, 2015). Opinions differ as to the main reason behind this pattern. According to Levinson (2016) children simply take longer to phrase their answers due to their less advanced language production abilities, however Pickering and Gambi (2018) suggest that the reason for the delay is that children's predictive abilities are behind that of adults'.

Consistent with Levinsos's proposal, a longitudinal study examining children's responses to questions found that the complexity of children's answers significantly predicted their response latencies, with longer conversational gaps preceding more complex answers (Casillas et al., 2016). Another study focused on the role of the question's predictability on the speed of turn-taking. In this study by Lindsay, Gambi and Rabagliati (2019), children answered more or less predictable questions while guiding a cartoon character through a maze in an interactive study. They showed that both 3-5 year old children and adults answered questions faster if the end of the question became predictable earlier in the sentence, suggesting that pre-school children are already capable of predicting the end of questions and can answer them faster if the prediction process can start sooner.

While the above studies provide crucial evidence regarding the factors influencing adults' and children's turn taking, and on what role prediction may play in the process, they do not examine the predictions made by participants directly, leaving the results open to alternative explanations that potentially exclude linguistic predictions. For instance, Pickering and Gambi (2018) suggest that people might answer utterances on the basis of a point at which the utterance could have potentially ended but did not (e.g. "What do you want to eat … for dinner tonight?") without actively predicting the ending of the sentence. If participants start preparing their answers as soon as they think they have enough information to do so, they might answer more predictable sentences faster even if no active prediction takes place. While the scope of the evidence resulting from turn-taking studies might be limited when it comes to assessing linguistic predictions, other experimental methods can provide more information on the subject.

1.3.3. Studies featuring the looking-while-listening paradigm

Unlike the studies discussed so far, studies using the looking-while-listening paradigm aims to assess predictions directly. In these studies, participants listen to sentences with highly predictable sentence endings (e.g. "The boy will eat the …") while their eye movements are followed as they scan an array of pictures depicting more or less fitting (i.e. predictable) endings (e.g. cake, ball, etc., see Figure 1.2, Altmann & Kamide, 1999). These studies have shown that adult participants tend to look at pictures depicting predictable sentence endings (e.g. picture of a cake) as opposed to unpredictable ones (e.g. picture of stone) long before the onset of the critical word (e.g. Altmann & Kamide, 1999; Borovsky, Elman & Fernald, 2012; Kamide, Altmann, Haywood, 2005).

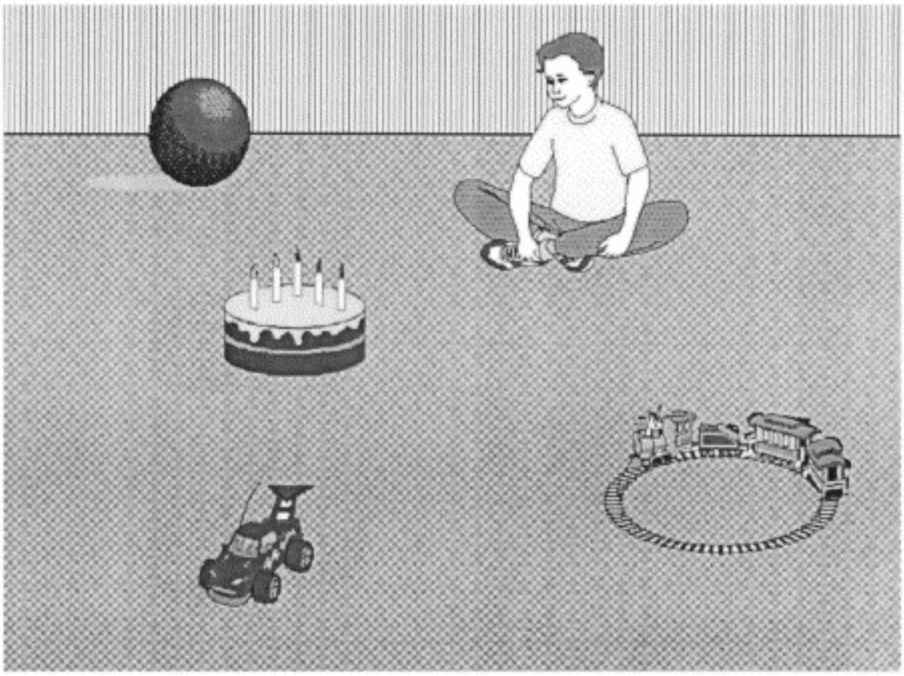

Figure 1.2. Picture array in a looking while listening paradigm

(Picture reproduced from Altmann & Kamide, 1999)

Results from other looking-while-listening studies suggest that information from a wide range of linguistic and non-linguistic sources influence adults' anticipatory eye movements. For instance, Altmann and Kamide (2007) showed that adults' looks towards potential final words are sensitive to verb tense: participants tended to look at an empty glass following "The man has drunk . . ." but at a full glass following "The man will drink....". Another study (Kamide, Altmann, & Haywood, 2005) found that the identity of the agent can also influence predictive looks (participants were more likely to look at the picture of a motorcycle after hearing “The man rides…” than after hearing “The girl rides…”).

Looking-while-listening studies with children have been similarly informative. A series of studies from Borovsky and colleagues suggest that children can use information from increasingly complex sources to guide their predictions as they get older (e.g. Borovsky, Elman, & Fernald, 2014; Borovsky, Sweeney, Elman & Fernald, 2014; Yazbeck, Kaschak & Borovsky, 2019). For instance, Borovsky et al. (2012) showed that 3 to 10 year old children can use the combination of agent and action cues to guide their anticipatory eye movements when listening to sentences such as “The pirate hides the … “. Participants from all three age groups (3 to 6 and 6 to 10 year old children and adults) looked more quickly at the picture of the treasure than at the picture of either the bone (action-related competitor) or the ship (agent-related competitor). A further study (Borovsky, et al., 2014) extended these findings to novel events. Children first learned about connections between agents, actions and objects (e.g. The monkey rides the bus) then took part in a looking-while listening experiment similar to the one described above. The authors found that while anticipatory looks of older (5-10 year old) children and adults were guided by their newly acquired knowledge (and looked at the picture of the bus after hearing “The monkey rides …”), younger children’s were not. Three - to 4 year-olds looked at the picture of the car instead of the bus in the previously discussed scenario. These results suggest a developmental trajectory where resource limitations might originally restrict what sources of information children are able to use for their predictions but over time, as more processing resources become available, children’s performance becomes more and more adultlike in these studies.

While the above results point to the existence of precise linguistic predictions that can take into account a wide range of information, other studies suggest that these looks can also be guided by spreading activation rather than anticipatory processing. For instance, a study by Kukona, Fang, Aicher, Chen, and Magnuson (2011) showed that upon hearing utterances such as “Bill will arrest ...”, adult participants looked at both the picture of the policeman as well as the robber, even though “the policeman” is not a highly likely sentence continuation (see also Kukona, Cho, Magnuson & Tabor, 2014). Interestingly, Gambi, Pickering and Rabagliati (2016) found a different pattern examining the anticipatory looks of 3 to 5 year old children. In their study, pre-schoolers only made anticipatory looks towards entities that are not only semantically but also syntactically predictable (e.g. they tend to look at the picture of a robber but not at the picture of the policeman after hearing sentences such as “Pingu will arrest the . . .”). This result further supports interpretations suggesting that children are capable of pre-activating linguistic input and shows they also use various sources of linguistic and word-related knowledge to make these predictions by the time they reach pre-school age at the latest. Other studies demonstrate that children can show adult-like performance in looking-while-listening studies at an even younger age: 2 year olds already tend to look longer at pictures of objects that would be a more predictable ending for the sentence before the onset of the target noun (Mani & Huettig, 2012; Mani & Daum & Huettig, 2016).

Due to this paradigm’s child-friendly nature, and the possibility that prediction plays a central role in language acquisition, several looking-while-listening studies have sought to target the relationship between prediction and learning. Some work has focuses on the potential connection between predictive abilities and language proficiency, but the results vary as to which (if any) aspects of proficiency seem to be connected to prediction. Mani and Huettig (2012) showed that toddlers’ prediction skills (measured using a version of the looking while listening paradigm) significantly correlated with their productive vocabulary, but not their receptive vocabulary. Borovsky and colleagues (2012) in turn found a correlation between 3 to 10 year old's predictive looks and receptive vocabulary, while Nation, Marshall & Altmann, (2003) found no correlation at all between 10 and 11 year olds reading skills and predictive skills (measured by the speed of their anticipatory looks).

Further studies have assessed the nature of children's linguistic predictions in order to examine whether they could form the basis of learning (e.g. Borovsky et al. 2012; Gambi, Pickering & Rabagliati, 2016; Lukyanenko & Fisher, 2016). For instance, Gambi and colleagues' (2016) study described earlier confirmed that children combine semantic associations (Borovsky et al., 2012) with syntactic knowledge (Huang, Zheng, Meng & Snedeker, 2013) suggesting that predictions made in childhood could form a viable basis for language acquisition.

In sum, the looking-while-listening paradigm has provided crucial information about the existence of predictions in adult- and childhood, about the sources of information that different age groups can benefit from when predicting, and, crucially, about prediction's relationship with language acquisition. However, the mechanisms behind what have been interpreted as predictive looks are controversial, and the limitations of looking-while-listening paradigm highlight some of the key outstanding questions in prediction research.

1.3.4. Prediction versus integration

One of the most pressing questions in prediction research is whether we are measuring the effects of prediction as opposed to the effects of integration. When prediction occurs, listeners activate the upcoming word prior to when it appears in the speech stream, which leads to facilitated processing when the word does appear. In contrast, when integrating, people are simply assessing how well a previously processed word fits in with the previous input. In other words, when measuring predictions we need to assess processes before the word of interest appears in the speech stream, whereas when measuring integration we are targeting processes after (or during) the point at which it was heard.

Rabagliati, Gambi and Pickering (2016) have suggested that what would be called "anticipatory looks" in looking-while-listening studies could be the result of integration and not prediction. It is possible that by the time participants heard "The boy eats the big …" , they have already looked at all the pictures showing potential sentence endings, and they only look at the picture of the cake the fastest because, after assessing how easy is it to integrate each picture's referent, they found 'cake' to be the most suitable sentence ending from those available as pictures. If this is the case, the differences in looking time would reflect differences in ease of integration and are therefore not informative about participants' predictions.

This is a particularly important issue in prediction research, as, for most studies, the critical measurements are taken during (or after) the processing of predictable versus unpredictable words, which makes it challenging to determine whether these differences are a result of the pre-activation of the more predictable word, or its easier integration. The difference between these two processes is crucial for error-based learning theories, as they propose that the error signal forming the basis of learning is the result of the computations based on predictive (and not integrative) processes. However, despite the importance of this distinction, differentiating between effects of prediction and integration also presents distinct challenges for other methods of investigating prediction, including, for instance, event related potential (ERP) studies measuring brain activity on the scalp while participants are attending to various stimuli.

1.3.5. Studies assessing event related potentials (ERPs)

Similarly to studies using the visual word paradigm, EEG studies measuring ERPs also tend to report processing differences between predictable and unexpected words. These studies uncovered differential neural responses to more or less surprising words, such as the final word of sentences 1.a and 1.b. These differences typically manifest as variability in the N400, a negative ERP component peaking 300 to 500 ms after word onset. This component has been shown to be sensitive to contextual predictability, with a roughly linear reduction in N400 amplitude as predictability increases (Kutas & Hillyard, 1984; Delong & Kutas, 2005). For example, Kutas and Hillyard (1984) found that the amplitude of the N400 component is the inverse function of the participants' expectations (measured by Cloze probability, the percentage of individuals that continue a sentence fragment with that item in an offline sentence completion task) of the sentences' final words, and suggested that the N400 might be a marker for linguistic predictions.

1.a. The day was breezy so the boy went outside to fly **a kite**.

1.b. The day was breezy so the boy went outside to fly **an airplane**.

However, studies such as the one described above suffer from similar constraints as those featuring the looking while listening paradigm: the effects of prediction and integration are hard to tease apart, as ERPs are measured after the onset of the critical word. Some EEG studies take a different approach, and instead measure processing differences before the onset of the critical word. A classic example of this approach is the 2005 study by DeLong, Urbach and Kutas. This study also assesses processing differences towards the end of the sentences, but with an inventive twist: they not only measure the processing differences on the predictable versus unpredictable final nouns ('airplane' versus 'kite' in sentences 1.a. and 1.b.) themselves, but also on the articles that precede them. This study has not only replicated the already established inverse cloze-probability-N400 relationship on the critical nouns, but also found an increased N400 on the article preceding the final noun if it did not match the article associated with the highly predictable ending (see Figure 1.3.). For instance, in sentence 1.a, the article 'an' proceeding the unexpected 'airplane' elicited a higher N400 than the article 'a' proceeding the expected 'kite'. As processing 'a' versus 'an' should only lead to different neural responses in this context if the next (yet unheard) word is also activated during article processing, this study has provided strong evidence that adults are not just integrating words as they appear in the speech stream, but that they are also capable of pre-activating upcoming words. While the replicability of the above results is currently debated (see Nieuwland et al. 2018; versus Yan, Kuperberg & Jaeger, 2017), analyses using different metrics of predictability than cloze values and alternative analytical approaches (e.g. Yan et al., 2017, Niewland et al., 2019; Delaney-Busch, Morgan, Lau & Kuperberg, 2019) have provided further evidence of the N400's sensitivity to the effects of linguistic pre-activation.

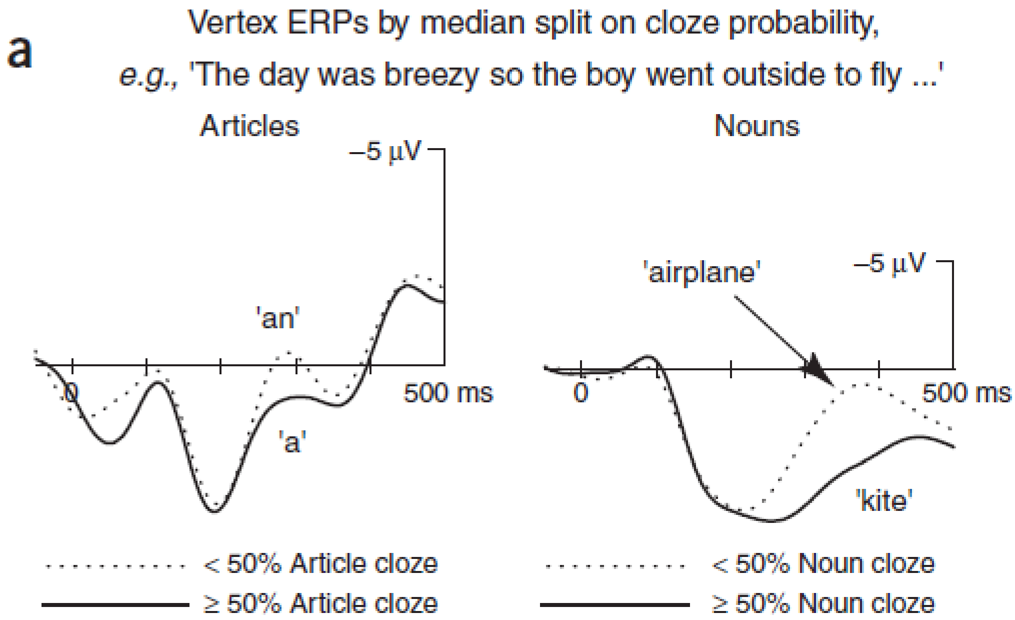

Figure 1.3. – N400 effects on predictable versus unpredictable sentence-final nouns and preceding articles (Figure reproduced from DeLong, Urbach & Kutas, 2005)

While the above studies assess how the immediately preceding linguistic context might lead to predictions, other EEG studies address how the wider linguistic environment affects predictive processes. For instance, Brothers, Dave, Hoversten, Traxler and Swaab (2019) compared the strength of participants' predictions (measured by N400 effects) based on the reliability of the speaker they were listening to. In this study, speakers were overall either reliable, and tended to complete sentences with words that were highly predictable, or unreliable, as their sentence completions were unpredictable (but plausible). Sentences from a reliable speaker showed larger facilitation effects with an earlier onset, suggesting that listeners engaged in enhanced anticipatory processing when a speaker's behaviour was more predictable. These results suggest that the strength of predictions can be regulated by the wider linguistic environment and are also in line with the possibility that predictive processing is not entirely automatic.

1.3.6. How often do people make predictions?

The above results highlight another important question in prediction research: how often do people make predictions? This question is of particular importance when considering error-based learning theories, as it is crucial to their operation that listeners make predictions for every upcoming word, as these predictions are seen as the basis of the proposed learning mechanism.

However, some accounts suggest that linguistic predictions may be optional, only occurring if sufficient resources are available (e.g. Mani & Huettig, 2016; Pickering & Gambi, 2018). Indeed, results from various studies show that predictions may be delayed or may not happen at all under certain circumstances, such as the input is noisier, the stimuli presentation is faster, or when there is interference from another task requiring high memory load (e.g. Huettig & Guerra, 2015; Brouwer, Mitterer & Huettig, 2013; Ito, Corley & Pickering, 2018). For example, Ito and colleagues (2018) found that adults' predictive eye-movements were delayed in a task in which they had to click on an object when it was mentioned while simultaneously preforming a working memory task, as opposed to when the clicking task was carried out alone.

Another set of studies examined the predictions of specific populations with lower language proficiency (such as L2 learners: Martin et al. 2013; Mitsugi & MacWhinney, 2016, older adults: Huang, Meyer, & Federmeier, 2012; Wlotko & Federmeier, 2012 or children: Borovsky et al. 2012; Mani & Huettig, 2012) and found that the members of the above groups are less likely to make predictions. Mani and Huettig (2012), for instance, reported that children with low vocabulary scores did not make predictive eye-movements towards potential sentence-final words when listening to highly predictable sentences.

Pickering and Gambi (2018) also suggest that, even when people make predictions, there might be a difference as to which features of the upcoming input they predict depending on the resources available. They propose that in earlier parts of utterances (when they had less time to formulate a prediction), participants may only predict the semantics (but not the phonology) of upcoming words.

It is also important to note that experimental studies that have provided evidence for linguistic predictions usually examine very specific scenarios. They typically make measurements on sentence-final words of highly predictable (versus unpredictable) sentences, which are presented with perfect articulation and at a slow speed, often while visual scaffolding is present (Mani & Huettig, 2016). While these studies may show us that participants can (and do) make predictions in circumstances that are ideal for predicting, they cannot provide us with any information regarding whether people are making predictions at all times, such as earlier in the sentence, at a faster speech rate or in less predictable environments. Therefore, despite the accumulating evidence that people can predict, we do not yet know whether they predict under all circumstances, and as such whether or not this crucial element of error-based learning theories is plausible.

In summary, targeting linguistic predictions is exceptionally challenging because experimental methods typically find it difficult to exclude alternative explanations of the results, which do not necessarily involve predictions. However, a growing body of results from a variety of methods, particularly from new approaches in neuroscience, suggests that adults – and probably children – are capable of pre-activating words in the speech stream ahead of time. Yet despite the convincing case for the existence of linguistic predictions, there are still some important open questions regarding linguistic predictions, including the one this dissertation seeks to address: are predictions used for language acquisition?

1.4. The prime surprisal paradigm

When addressing prediction's role in language acquisition, the most crucial limitation of the research discussed above is that it primarily concentrates on whether or not people are capable of making linguistic predictions, and if so, what influences these predictions. The majority of prediction research (with the exception of studies assessing turn-taking) does not address how predictions affect future linguistic behaviour, and most importantly does not examine whether the success or failure of predictions actually changes linguistic knowledge. In other words, the research discussed above cannot determine whether predictions form part of an error-based learning mechanism.

In order to address prediction's role in language acquisition, in this project, we used the prime surprisal paradigm, a method targeting the influence of the predictability of previous linguistic input on subsequent language behaviour. This is an ideal tool for assessing the whole of the error-based mechanism, as instead of focusing exclusively on the existence of predictions, it also assesses how exposure to sentences that are likely to lead to either successful or erroneous predictions influences subsequent linguistic behaviour.

1.4.1. The structural priming paradigm

The prime surprisal method is rooted in the structural priming paradigm, which in turn is based on the observation that speakers tend to repeat linguistic structures they have recently encountered. In production priming studies, participants are typically exposed to a (written or spoken) prime sentence and are then expected to produce a target sentence. If the prime structure is re-used in the target sentence at a greater frequency than after a prime featuring an alternative structure, it is then taken as support of the existence of abstract representations in the given population (see Mahowald, James, Futrell & Gibson, 2016 for a meta-analysis on structural priming in language production). In an early production priming study, for instance, Bock (1986) examined whether participants tend to repeat previously processed structures in dative and transitive sentences while describing pictures. This study demonstrated significant priming effects in sentences featuring both structures: adults produced 23% more double object dative (DOD) targets after DOD rather than PD (prepositional dative) primes, and 8% more descriptions featuring a passive structure after a passive rather than active transitive prime. Further studies have suggested that these effects were not simply driven by similar thematic roles, closed class function words or prosodic patterns (Bock & Loebell 1990; Bock, 1989), providing further evidence that the effects are syntactic in nature. However a more recent study by Ziegler, Bencini, Goldberg and Snedeker (2019) casts doubts on the exclusively syntactic nature of structural priming. Ziegler and collegueas showed that while passives and intransitive by-locatives (both containing the by-preposition) prime passives similarly, intransitive locatives (not featuring *by*) do not prime passives. These results suggest that passive priming is dependent on the presence of the preposition *by*, thus the overlap of abstract syntactic structure alone may not be sufficient for passive priming.

Priming studies typically feature syntactic structures that can alternate between different syntactic forms that carry similar meanings, like the dative or the transitive structures seen in Bock's (1986) study. Datives, for instance, alternate between prepositional object datives (PD, e.g. "The king sold a rabbit to the queen.") and double object datives (DOD, e.g. "The king sold the queen a rabbit."). Similarly, transitive sentences alternate between an active (e.g. "The king scared the queen.") and a passive (e.g. "The queen was scared by the king.").

Structural priming has long been used to examine the nature of abstract structural representation, and more recent studies have found structural priming effects under various conditions, such as in comprehension as well as production (e.g. Branigan, Pickering & McLean, 2005; Traxler, Tooley & Pickering, 2014), in different linguistic modalities (e.g., Branigan, Pickering & Cleland, 1999; Pickering & Branigan, 1998), in languages other than English (Cai, Pickering & Branigan, 2012), and even across languages (e.g. Kantola & van Gompel, 2011). Most importantly for our purposes, structural priming effects have also been observed in young age groups, including children (e.g. Savage, Lieven, Theakston & Tomasello, 2003; Messenger, Branigan, McLean & Sorace, 2012; Peter et al., 2015).

1.4.2. Developmental structural priming studies

As the structural priming paradigm is thought to tap into the existence of abstract linguistic representations, it has been used as a major tool for examining when these representations develop in childhood. By comparing priming effects between prime and target pairs sharing different levels of lexical and syntactic content, this method has been used to assess the predictions of the early abstraction and lexical constructivist accounts of language acquisition (e.g. Hirsh-Pasek & Golinkoff, 1996; Olguin & Tomasello, 1993). The first account argues that children have abstract syntactic representations from a young age (and predicts abstract priming effects from early on), while the latter account proposes that abstract syntactic structures develop from lexically dependent structures (and predicts that children will initially only show priming effects when there is lexical overlap between prime and target items).

Children at an early age already showed priming effects in structural priming studies. For instance, Bencini and Valian (2008), found that three-year olds were more likely to produce passive sentences after a passive (than an active) prime. However, existing results are mixed regarding whether these effects are stronger in the presence of lexical repetition (i.e. whether a lexical boost is present or not; see Savage et al. 2003 versus Rowland, Chang, Ambridge, Pine & Lieven, 2012). While targeting the lexical boost has not lead to a conclusion regarding the early abstractionist versus lexical constructivist debate, studies finding abstract structural priming without lexical overlap at an early age (e.g. Bencini & Valian, 2008; Messenger, et al., 2012) have provided support for the early abstraction accounts.

However, while these studies have revealed a great deal regarding how children's syntactic knowledge is stored at any given point in development, they do not aim to describe how children progress from one knowledge state to another, and in themselves do not offer evidence about language acquisition mechanisms.

1.4.3. Prime surprisal

The prime surprisal method builds on classic priming studies, but focuses on whether more surprising (less predictable) sentences result in stronger priming effects than predictable ones.

This paradigm builds on the existence of different verb biases: dative verbs tend to appear more often in one of the dative variants (e.g. DOD) than in the alternative one (e.g. PD). For instance, the verb *sell* is PD-biased and appears more often in PD than DOD sentences, while the verb *give* prefers the DOD structure over PD. Consequently, a sentence featuring *give* is more predictable in a DOD structure (e.g. *The king gave the queen an apple*), than in a PD structure (e.g. *The king gave an apple to the queen*).

Error-based learning theories propose that as listeners are constantly predicting the following word, when they reach the verb (*give*) in the above sentence they will start anticipating a DOD structure (*The king gave … the queen an apple*). If they end up hearing the PD structure instead (*The king gave … an apple to the queen*), an error-signal is produced based on the difference between the expected and actual input. In contrast, if the expected DOD structure is processed, no such signal is produced. This error signal is than used to adjust the weights supporting the relevant aspects of syntactic knowledge (in this case frequencies associated with the DOD structure). This adjustment then initiates two important processes. First, and crucially for the experimental assessment of error-based learning theories, it leads to an increased likelihood of immediate structure repetition, in other words to a prime surprisal effect. Second, and most importantly for the learning aspect of these theories, these weight changes accumulate over time and lead children closer to adult-like language with each interaction. According to the Dual Path model, these processes continue into adulthood, but in adulthood they facilitate the accommodation to different speakers or linguistic scenarios rather than language acquisition per se.

1.4.4. Adult prime surprisal studies

Previous studies have found increased priming effects after surprising as opposed to predictable prime sentences; in other words, they showed prime surprisal. The first demonstration of this effect comes from Jaeger and Snider's (2008) work, in which they reanalysed dative structures collected from a corpus of spontaneous speech (Bresnan, Cueni, Nikitina & Baayen, 2007). They found stronger priming effects after PD sentences that included mis-matched (DOD biased) verbs as opposed to those featuring matching (PD biased) ones. For instance, the verb 'cost' is heavily biased towards the DOD structure (see 2.a. versus 2.b.) and indeed PD sentences including this, or other DOD biased verbs (such as 2.a.), led to larger priming effects than those PD sentences that included PD biased verbs. However, this effect only appeared in sentences including PD structures and not in those featuring DODs. The lack of prime surprisal in DOD sentences might either be due to the inverse frequency effect (showing that less frequent structures cause less strong priming) or to the lower variation in the verb's sub-categorisation bias in this part of the dataset.

2 a. A hard disk drive would cost several thousand dollars to the consumer...

2 b. ...inaccurate credit information could cost the consumer tens-of-thousands of dollars...

Further studies have expanded the scope of prime surprisal effects in adults. Showing that prime surprisal is not unique to English, similar results have been established in a study on Dutch datives by Bernolet and Hartsuiker (2010) who found stronger priming effects after DOD primes including structures when they were paired with mis-matching (PD biased) verbs. Fine and Jaeger (2013) have expanded prime surprisal effects to comprehension as well in their work featuring the re-analyses of Thothathiri and Snedeker's (2008) dataset. They showed that more surprising prime structures (including mis-matching as opposed to matching verb-structure pairings) led to stronger expectations that the prime structure will be repeated in the target sentence. In a set of three studies, Jaeger and Snider (2013) also showed that expectations are not only sensitive to prior experience (verb biases) but also recent experience (accumulated over the course of the study), showing how error-based learning could facilitate adapting to the varying linguistic qualities of different linguistic situations.

1.4.5. Prime surprisal in children

Importantly for our purposes, prime surprisal effects have also been demonstrated in a study involving child participants. Peter, Chang, Pine, Blything and Rowland (2015) carried out a prime surprisal study including three age groups: 3-4 and 5-6 year old children and adults. To make the study suitable for child participants, Peter and colleagues embedded the priming study in a bingo game in which participants collect cards on bingo boards while taking turns at describing video animations involving familiar figures, either named characters from age-appropriate popular culture, such as Dora and Boots, or generic pairs, such as king and queen (see also Rowland et al. 2012 for another developmental priming study using a similar set-up). The study involved dative sentences, such as those in examples 3.a. and 3.b..

3.a. Dora gave a bunny to Boots.

3.b. Wendy gave Boots a bunny.

The goal of the study was to investigate when and how children develop adult-like verb structure links. The effect of three parameters was assessed: verb overlap between prime and target, target verb bias, and prime verb bias. This study has replicated some previous findings demonstrating priming effects in all three age groups and a lexical boost (increased priming when there is verb overlap between prime and target sentences). Critically, they found that both the bias of the target verb and crucially the prime verb had an effect on structure choice in the target sentences, showing that children are already sensitive to verb-biases from a pre-school age. They also found that prime surprisal effects (larger priming after mis-matching prime verb and structure pairing) were increased in the younger age groups, with 3-4 year olds showing the largest and adults showing the smallest (only marginally significant) effects. Both of these results are in line with the predictions of error-based learning theories.

The Dual-path model suggests that these results arise due to the error-based learning mechanism, according to which participants in this study are constantly predicting the next upcoming word. When they hear a verb ('give'), they predict the next word based on the verb's bias. For instance, for a video depicting Dora giving a bunny to Boots, after hearing the DOD-biased verb 'give', participants would be most likely to predict the first word that would appear in a DOD structure: 'Boots' (see example 3.a.). If instead they end up hearing the first word of the unpredicted PD structure (bunny) (see 3.b.), they detect a discrepancy. This in turn results in the production of an error-signal leading to not only a change in the weights supporting syntactic knowledge, but also to a higher likelihood of the repetition of the previously processed structure (an increased priming effect). However, if the participant hears the DOD structure after a DOD biased verb (see sentence 3.b.), they are likely to have made a correct structure prediction, and as there is no discrepancy between the predicted and the actual input, no error signal is produced and neither the supporting weights nor the likelihood of a priming effect are changed.

The stronger prime surprisal effects in the younger age groups are also predicted by the Dual-path model, as the model proposes that as more linguistic experience is accumulated, syntactic representations become stronger and less malleable, and these stronger representations experience a smaller shift (leading to reduced prime surprisal effects) after each linguistic interaction.

As seen above, the prime surprisal method is an effective tool for targeting error-based learning, as it can assess all of the proposed learning mechanisms, from predictions themselves to how their outcome influences subsequent language use. It can also be used to design studies that are suitable for younger age groups. Although the above qualities make it an excellent tool for discovering more about error-based learning, to our knowledge only one published study has used this method with children, while the number of adult studies (discussed in section 4.4. Adult prime surprisal studies) is also limited. In the current project we combined this paradigm with other experimental techniques to learn about prediction's role in language acquisition and assess error-based learning. The following three experimental chapters will be discussing three sets of studies addressing the following research questions:

Chapter 2 – Do prime surprisal effects appear in sentences featuring structures other than datives?

Chapter 3 - Do children (and adults) learn more from unexpected structures than expected ones?

Chapter 4 – Do adults process more and less expected sentences differently? (ERP evidence)

The following chapter will then discuss two prime surprisal studies including adult participants, one of which is a replication of previous prime surprisal studies featuring dative structures and another which aims to expand the examination of prime surprisal to a new structure: transitive sentences.

CHAPTER TWO: PRIME SURPRISAL IN DATIVE AND TRANSITIVE STRUCTURES

2.1. Introduction

Prediction, the ability to activate upcoming words before we encounter them, plays a crucial role in multiple theories of language processing. It might contribute to successful dialogue (Pickering & Garrod, 2013), help us keep up to date with our ever-changing linguistic environment (Jaeger & Snider, 2013) and it may also be vital for language acquisition (Chang, Dell & Bock, 2006).

Several predictive accounts of language processing operate via some form of an error-based learning mechanism (e.g. Chang, Dell & Bock, 2006; Ramscar, Dye & McCauley, 2013). These mechanisms propose that people constantly predict upcoming linguistic input, evaluate these predictions and update their knowledge based on whether the predictions were correct. In the Dual-path model (Chang, Dell & Bock, 2006), for instance, the error-based learning mechanism is embedded in a connectionist, frequency-based model that targets the development of syntax. This model proposes continuity between processing in child- and adulthood: the error-based learning mechanism underlying syntactic acquisition in childhood is the same mechanism that leads to syntactic alignment in adulthood. This learning mechanism compares the predicted and actual linguistic input and generates an error signal if it detects a discrepancy. The error-signal is than used to adjust the weights supporting syntactic knowledge in all age-groups. During this process, the listener's knowledge further approximates the linguistic environment with every interaction: for adults this facilitates accommodating to different linguistic situations, conversational partners or registers, and the same process leads children closer to adult-like language with each linguistic exchange.

Despite the prominent role given to prediction in multiple accounts of language acquisition, there are still several open questions about the key characteristics of linguistic prediction. Opinions differ, for instance, on how many consecutive words people can predict, which linguistic levels are predicted, and whether predictions are all-or-none or if multiple alternative predictions are considered in parallel (Kuperberg & Jaeger, 2016). The fact that these questions remain open is partly due to the research methods that are typically used to evaluate linguistic predictions. Most studies measuring predictions assess whether a word that is more predictable given the preceding context is processed differently than one that is less predictable. For instance, adults read more predictable words faster than surprising ones (e.g. Schwanenflugel & Shoben, 1985) and they also look at pictures faster if they are depicting a more predictable referent (e.g. Altmann & Kamide, 1999). These results could demonstrate that participants have pre-activated the predictable words (but not the surprising ones), which made it easier to process them once they appeared in the input. However, it is also possible that these words were not pre-activated, but rather that they required less effort to process as they were easier to integrate with the preceding context (e.g. Rabagliati, Gambi & Pickering, 2016). While the source of the above results is debated, more recent studies using on-line experimental techniques (e.g. Lau, Holcomb & Kuperberg, 2013) and novel analytical approaches (e.g. Niewland et al., 2019) provide clearer evidence for linguistic predictions.

However, from the perspective of this paper, the above research still has a crucial limitation: its main aim is to evaluate the existence of people's predictions. It does not assess a major element of proposed error-based learning mechanisms – how predictions actually affect subsequent language use. In our study we use a different method, prime surprisal. This method provides a more complete picture of error-based learning mechanisms, as it targets how processing predictable and surprising input influences subsequent language use.

Prime surprisal builds on the structural priming paradigm, which is typically used to examine syntactic representations by assessing whether people repeat previously processed linguistic structures (e.g. Bock, 1986). These studies usually feature syntactic structures that alternate between different forms that have similar meanings, such as the dative or transitive structures. Datives, for instance, can appear either as prepositional datives (PD, e.g. *The king sold a rabbit to the queen.*) or as double object datives (DOD, e.g. *The king sold the queen a rabbit.*). Similarly, transitive sentences can either be expressed using an active (e.g. *The king scared the queen.*) or a passive (e.g. *The queen was scared by the king.*) structure.

The prime surprisal method takes classic priming studies a step further and examines whether more surprising primes lead to stronger priming effects, as predicted by error-based learning theories. This paradigm relies on the fact that the likelihood of verbs appearing in alternative sentence structures is different for each verb. For instance, while overall DODs appear more often in adult language use, all verbs have their own specific preferences: while the verb *bring* occurs more often in a PD structure than in a DOD structure, the verb *give* prefers the DOD structure. Consequently, a sentence including the verb *bring* is more surprising in a DOD structure (e.g. *The king brought the queen an apple.*) than in a PD structure (e.g. *The king brought an apple to the queen.*). Error-based learning accounts predict stronger priming effects (resulting from bigger change in syntactic representations) after surprising (e.g. PD-biased verb in a DOD structure) than predictable (PD-biased verb in a PD structure) primes.

Previous studies have indeed found increased priming when structures appeared with mismatching as opposed to matching verbs (e.g. Fine & Jaeger, 2013). Prime surprisal effects have been found in production (Jaeger & Snider, 2013), comprehension (Fine & Jaeger, 2013) as well as in a language other than English (e.g. Dutch; Bernolet & Hartsuiker, 2010). Prime surprisal has even proven to be sensitive to recent experience accumulated over the course of the study as well as prior experience (verb-biases), highlighting error-based learning's possible role in rapid adaptation to changeable linguistic scenarios (Jaeger & Snider, 2013).

Prime surprisal effects based on prior expectations have also been demonstrated in a study including child participants. Peter and colleagues (2015) carried out a dative prime surprisal study embedded in a bingo game with three age groups (3-4 and 5-6-year-old children and adults). They found both structural priming and prime surprisal effects in all age groups; furthermore, the prime surprisal effects were stronger in the younger groups. The error-based theories predict that prime surprisal effects shrink with age because syntactic representations become stronger and less malleable as more linguistic experience is accumulated, and these stronger representations experience a smaller shift (resulting in smaller prime surprisal effects) after each linguistic interaction.

The dual path model attributes the above results to the workings of the error-based learning mechanism. It suggests that participants constantly predict the next upcoming word and, as they hear a verb (e.g. *bring*), they predict the first words of the structure that most often follows this verb (e.g. *bring* -> ***the*** *ball to Bob* (PD structure)). If they end up hearing the first words of the alternative structure (e.g. *bring* -> ***Bob*** *the ball* (DOD structure)) they detect a discrepancy. An error-signal is produced which then leads to both a change in the weights supporting syntax and to a higher likelihood of the participant reproducing the structure that they have just heard (i.e. a larger priming effect). However, if after hearing the PD-biased verb *bring* they hear the matching PD structure, their prediction is more likely to have been correct. In the absence of an error-signal, the weights supporting syntactic knowledge are not affected and the likelihood of priming is not increased.

While prime surprisal studies can provide useful insight into error-based learning, we can only draw limited conclusions from existing studies because, to our knowledge, prime surprisal studies have to date almost exclusively featured dative structures. Datives provide an excellent test case for examining error-based learning: several verbs have strong dative verb biases and verbs appear early in sentences. The verbs' biases can than then set up strong expectations for the subsequent dative structure at the beginning of the sentences. It can then be effectively contrasted how the dative structures that follow the verb influence subsequent language production depending on whether they matched (predictable) or mis-matched (surprising) the verb's bias. However, according to the Dual-path model structures other than datives should also be sensitive to prime surprisal. Verbs are less strongly biased in other structures, for instance in active versus passive sentences. Furthermore the verbs location in the sentence is also different. It may already become clear to the listener whether they are encountering an active or a passive sentence when the verb first appears. Examining whether these structures are also sensitive to prime surprisal would be informative about the different conditions error-based learning mechanisms might operate in.

To learn more about linguistic predictions and about how they participate in error-based learning, we carried out two prime surprisal studies featuring both dative and transitive sentences. We used a design that has led to both priming (Rowland et al., 2012) and prime surprisal (Peter et al., 2015) effects with dative structures previously to provide a good baseline for comparison for our transitive results. While based on previous literature we expect to replicate structural priming effects with both the dative and transitive structures and detect prime surprisal effects in the dative study, we do not yet know whether prime surprisal also occurs with transitive structures.

2.2. Methods

The goal of this study was to learn more about linguistic predictions and their role in error-based learning mechanisms. We carried out two prime surprisal studies: one featuring a structure that has shown to be sensitive to prime surprisal in previous studies: the dative structure, and a second study featuring a structure not yet included in published prime surprisal studies, the transitive structure.

2.2.1. Participants

One hundred and twenty adults from the University of Liverpool student participation pool took part in the study and received course credit for their participation. All participants took part in both the dative and the transitive study, in one of two counterbalance conditions (dative or transitive first). Four participants were excluded from both of the studies as they produced ‘other’ responses for more than half of the target trials in one of the studies. Exclusion criteria for the target sentences will be discussed in the 2.2.6. Coding section.

2.2.2. Design

Both studies used a 2x2 within-subjects design, where the variables were prime type (DOD vs. PD in the dative and active vs. passive in the transitive study) and verb-bias match (match or mismatch). The dependent variable was the choice of dative or transitive structure in the target trials.

2.2.3. Visual stimuli

The study featured images created in GNU Image Manipulation Program (The GIMP Team, 1997-2014), the images were presented in E-prime 2.0 software (Psychology Software Tools, Inc., 2000). Both studies contained 80 images overall: 20 images for the prime sentences, 20 images for the target sentences and 40 images for the filler sentences. The images depicted the action featured in the sentences (e.g. *give, scare*) using the characters in the sentence (e.g. *Lisa* and *Bart*), together with a written version of the verb above the picture (see Figure 2.1.).

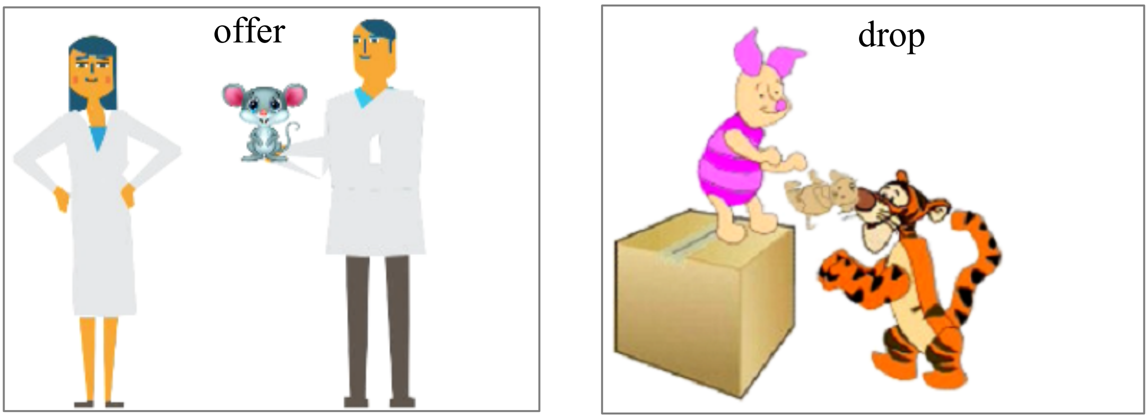

Figure 2.1. - Example prime and target pictures from the dative study

Previous priming studies using a similar set-up to the current one (Peter et al., 2015; Rowland et al., 2012) used a stem competition technique for the target sentences to ensure the participants use the correct target word. When using the stem competition technique, the experimenter provides the sentence stem (e.g. *Bart gave* ...) and the participant repeats the stem and the finishes the sentence (e.g. *Bart gave Lisa a fox*.). However, as in transitive sentences the structure decision point (where it becomes clear whether an active or passive sentence is presented) is before the verb (e.g. ***Lisa scared*** *Bart.* versus ***Bart was scared*** *by Lisa*.) it was impossible to use this technique while also giving participants a free choice between the two alternative transitive structures. Therefore, to ensure the use of the correct target verb, we included a written version of the verbs to be used above each prime and target picture.

The images included 10 pairs of donor and recipient characters. Half of them were cartoon characters with proper noun names: *Tigger and Piglet*, *Dora (the Explorer) and Boots*, *Marge and Homer*, *Lisa and Bart* and *Bob (the Builder) and Wendy*. The other characters were referred to with determiner and noun NPs: *the prince and the princess*, *the king and the queen*, *the student and the teacher*, *the doctor and the nurse* and *the boy and the girl*. Particular donor and recipient characters were always featured together. A further 10 animate characters were featured as objects and were referred to with non-definite determiner and noun NPs: *a monkey*, *a cat*, *a fish*, *a butterfly*, *a fox*, *a bird, a mouse*, *a bunny*, *a baby*, and *a bird*. These characters were consistently paired with one pair of characters (e.g. the *bunny* was always be featured with *Bob and Wendy*). In order to keep the intonational contour of the prime and target sentences different we always featured characters with proper noun names in the target sentences if the primes contained determiner and noun NP-s and vice versa (see Rowland et al., 2012).

Prime images were always paired with target images including different characters from those in the prime. In order to control for the possibility that direction of transfer may influence structure choice, the images depicted the direction of motion of transfer actions equally often from right-to-left and from left-to-right.

2.2.4. Sentence stimuli

The dative and the transitive study contained 80 sentences each. The dative study contained 20 dative prime and target pairs and the transitive study featured 20 transitive prime-target pairs. Both studies featured 40 filler sentences. The prime and target verbs featured in both studies were selected based on their verb-structure frequencies in the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001, available on CHILDES, MacWhinney, 2000) as this was used as a basis of Peter and colleagues' (2015) prime surprisal study demonstrating dative prime surprisal effects with adult participants. The Manchester corpus consists of transcripts of audio recordings from a longitudinal study of 12 monolingual English-speaking children between the ages of approximately 2 and 3 years who grew up in major English cities.

In the dative study, the prime sentences appeared half the time as DOD (double object dative) sentences and half the time as PD (prepositional dative) sentences. Half of the verbs featured in the prime sentences were DOD-biased while the other half were PD-biased, which resulted in four prime categories featuring five sentences each: (1) DOD structure – DOD-biased verb (2) DOD structure PD – biased verb (3) PD structure – DOD-biased verb and (4) PD structure PD – biased verb. The target sentences were produced by the participant based on the target picture (as either DOD or PD sentences). The target sentences contained equibalanced verbs (which appear similarly often in both DOD and PD structures). The filler pictures depicted non-causal actions that can be described with intransitive sentences (e.g. *The boy was smiling*.) in both studies.

In the transitive study, prime sentences appeared half the time as passive sentences and half the time as active sentences. The Manchester corpus did not contain enough overall passive-biased verbs. Thus the featured verbs are overall active-biased, but half of the verbs featured in the prime sentences are more passive-biased (appeared in the passive structure more often in the child-directed adult speech in the Manchester corpus than the verbs in the other bias group) while the other half are less passive-biased (appeared in the passive structure less often in the Manchester corpus, than the verbs in the other bias group) resulting in four prime sentence categories featuring 5 sentences each: (1) passive structure – more passive-biased verb, (2) passive structure – less passive-biased verb, (3) active structure – more passive-biased verb, (4) active structure – less passive-biased verb. Similarly to the dative study, the target sentences were produced by the participants based on the target images (as either active or passive sentences) and they contained the most equibiased verbs in the corpus (i.e. that appeared less often as passives than the more passive-biased prime verbs but more often than the less passive-biased verbs). The verbs feartured in both studies (together with their DOD frequency based on the Manchester corpus and their classification as prime or target in this study) are featured in Table 2.1.. Semi-randomised stimuli lists were created in which the prime and target sentences always followed each other and between each prime-and target pair there was always a filler-filler sentence pair.

2.2.5. Procedure

The procedure was identical in the dative and the transitive studies. Both studies used the bingo paradigm adapted from Peter and colleagues' (2015) prime surprisal study (see also Rowland et al 2012). The priming studies were embedded in a bingo game in which the experimenter and the participant took turns to describe images that appeared on a laptop screen.

The sessions started with the experimenter introducing all the characters involved in the tasks by showing the participants pictures of them before the bingo game started. A short practice bingo session was also carried out to ensure that the participants understood the task. The practice session included intransitive sentences featuring three characters each (e.g. *The king and the queen were playing with the* *cat.*).

During the session, the experimenter and the participant set in front of the computer side by side. First a picture appeared on the left side of the screen, where the experimenter was sitting. The experimenter than described this picture (using the appropriate prime sentence) than asked the participant to repeat the prime. Then another picture appeared on the participant's side of the screen and the participant then produced a target sentence based on this picture. Each dative or transitive prime–target pair was immediately followed by an intransitive filler-filler pair.

After each prime-target pair, a question mark appeared on the screen followed by either a smiley or frowny face to indicate whether or not a bingo card is available. The first person to fill the bingo grid with bingo cards is the winner of the game and the experiment is designed so that the participant will always win.

2.2.6. Coding

The experiments were audiotaped, and the target utterances were coded off-line. In the dative study, a target response was considered a DOD if it contained the correct target verb followed by two noun phrases (e.g. *Lisa gave Bart a cat.*) , and a PD if it contains the correct target verb followed by a noun phrase and a prepositional phrase headed by *to* (e.g. *Lisa gave a cat to Bart.*). Responses were coded as 'other' if (a) the participant failed to repeat the prime correctly (even after help), (b) if the participant produced a non-target verb, or (c) if the target sentence could not be classified as a DOD or PD response based on the above criteria (e.g. target responses containing a preposition other than *to* or incomplete datives such as *Dora threw the ball at Boots.* or *The king gave the sandwich*.).

In the transitive study a target response was considered a passive if it contained the correct verb preceded by the patient and *was* and followed by a *by* phrase and the agent (e.g. *Lisa was scared by Bart.*). A target response was considered an active if it contained the correct target verb preceded by the agent and followed by the patient (e.g. *Bart scared Lisa.*). Responses were be coded as 'other' if (a) the participant failed to repeat the prime correctly (even after help), (b) if the participant produced a non-target verb, or (c) if the target sentence cannot be classified as an active or passive response based on the above criteria (e.g. target responses containing a preposition other than '*by*' such as *Bart was infuriated with Lisa*. or short passives, such as *Bart was infuriated.*).

2.3. Results

The aim of this work was to replicate prime surprisal effects in dative structures as seen in previous studies (e.g. Jaeger & Snider, 2013; Peter et al., 2015) and examine whether these effects also appear with transitive structures.

We analysed the dative and transitive datasets separately by fitting maximal logistic mixed effect models (Baayen, Davidson & Bates, 2008; Jaeger, 2008) using the lme4 package in R (R core team, 2012, version 3.6.1). We carried out model comparisons to obtain likelihood ratios and p-values on maximal models where the random effect structure was only simplified if the model did not converge (Barr, Levy, Scheepers & Tily, 2013). The dependent measure was be the production of either of the DOD structures (DOD=1, PD=0) or of the transitive structures (passive=1, active=0). All factors were effect/sum coded (Wendorf, 2004).

2.3.1. Results from the dative study

Figure 2.2. shows the mean proportion of DOD responses after PD and DOD primes involving either structure matching or mis-matching verbs. These analyses tested whether structural priming (a larger proportion of DOD responses after DOD

rather than PD primes) and immediate prime surprisal effects (larger priming effects if the prime verb's bias did not match the prime structure) were present in this dataset.

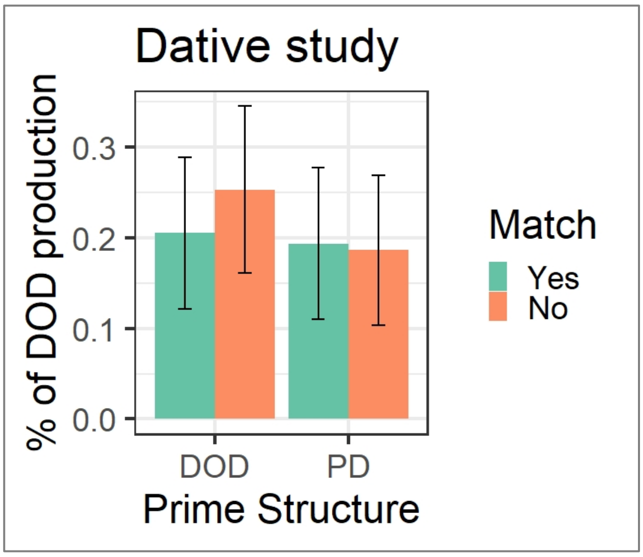

Figure 2.2. – Percentage of DOD responses after matching and mis-matching PD and DOD primes. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

Our model included (a) prime structure (DOD or PD prime) and (b) prime-bias match (depending on whether the prime structure matched the prime verb bias)[1] as fixed effects and by-participant and by-item random slopes. The results revealed a numerical structural priming effect, in that participants produced 3.90% more DOD structures after DOD than after PD primes, but this effect failed to reach significance, β= 0.11, $\chi 2$ (1) = 3.35, p = .07. While participants were 4.50% more likely to repeat the prime structure in the target sentences after mis-matching than after matching sentences (which is the premise of prime surprisal), neither the effect of prime bias match, β= 0.07, $\chi 2$ (1) = 1.21, p = .27, nor its interaction with prime type reached

[1] We also conducted this analyses with continuous verb-bias predictors (based on DOD-frequencies based on both the Manchester corpus and our norming study, see Table 2.1.) where the interaction of prime type and verb-bias would have indicated a prime surprisal effect. None of these analyses changed the pattern of results, neither the main effect of verb-bias, nor its interaction was significant in any of these analyses (all ps > .32).

significance, β= -0.56, $\chi2$ (1) = 1.41, p = .23.

As previous prime surprisal studies (e.g. Jaeger & Snider, 2008) found that the prime surprisal effects only occurred after prime sentences featuring the less frequent prime structure, we also carried out two sets of analyses on the target structures following the DOD or PD primes separately. The first of these analyses only included the target sentences following DOD primes. Note that in other studies, overall, PD structures are less frequent than DOD structures. In the current dataset however the majority of answers were PDs: overall 79.89% of the target sentences contained a PD structure), therefore we could expect a prime surprisal effect in the (less frequent) DOD-prime part of the dataset. The first model in this analyses included prime bias match as a fixed effect and by-participant random slopes for bias match and by-item random slopes. The results revealed a numerical but non-significant prime surprisal effect, where participants were 4.3% more likely to repeat the DOD structure after a mis-matching than after a matching prime (β= -0.54, $\chi2$ (1) = 2.33, p = .13).

The second set of analyses was carried out on the target sentences following PD primes. The first model in this analysis included prime bias match as a fixed effect and by-participant random slopes for bias match and by-item random slopes. The results revealed no effect of bias match (β= -0.04, $\chi2$ (1) = 0.09, p = .76) providing no evidence for a prime surprisal effect after PD primes.

In summary, in the dative prime surprisal study, there was a numerical, but non-significant priming effect and a numerical, but non-significant effect of prime surprisal in the less frequent (DOD) structure. The reason why we emphasise that there were numerical, albeit non-significant effects, is because the size of the structural priming effect (3.90%) and the prime surprisal effect in the DOD structures (4.30%) was numerically as large as, or larger than, the significant structural priming effects in the literature (which tend to be between 3.00 and 4.00%; see Mahowald et al, 2016). Thus we are reluctant to simply dismiss these data as showing null effects. We return to this point in the discussion.

2.3.2. Results from the transitive study

Figure 2.3. shows the mean proportion of passive responses after passive and active primes involving either matching or mis-matching verbs. The goal was to determine whether structural priming (a larger proportion of passive responses after passive than active primes) and immediate prime surprisal effects (larger priming effects if the prime verb's bias did not match the prime structure) are present in the dataset.

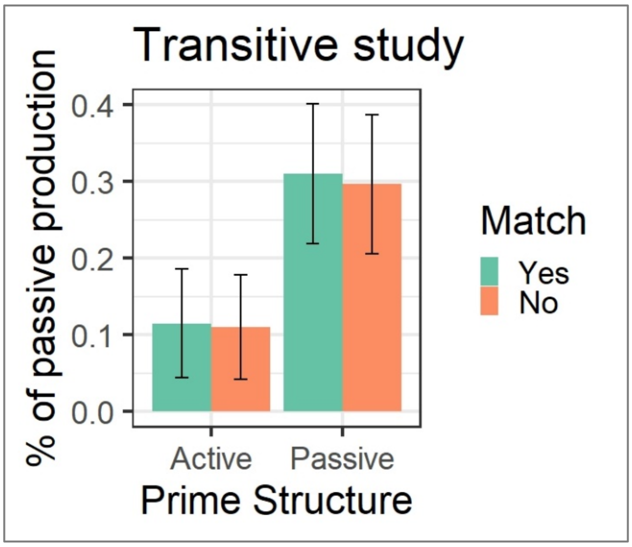

Figure 2.3. - Percentage of passive responses after matching and mis-matching passive and active primes. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

The first model included (a) prime structure (passive or active prime) and (b) prime-bias match (matching or mis-matching prime verb bias and prime structure)[2] as fixed effects and by-participant random slopes for prime and by-item random slopes. The results in this study revealed a significant structural priming effect, participants produced 19.1% more passive structures after passive than after active primes (β= -0.69, $\chi2$ (1) = 118.12, p < .0001). However, contrary to our prime surprisal prediction, participants were numerically less likely to repeat the prime structure in the target sentences after mis-matching than after matching sentences (1.30% less likely) and neither the effect of prime bias match, β= -0.032, $\chi2$ (1) = 0.06, p = .81, nor its interaction with prime, β= 0.001, $\chi2$ (1) = 0.001, p = .99, had a significant effect, indicating that prime surprisal effect were not present in this dataset.

As with the dative dataset we also examined the possibility that the prime surprisal effect only appears after one of the prime structures by examining the target structures following the passive and active primes separately. We first tested the effects following passive primes, and included bias-match as a fixed effect and by-subject and by-item random slopes. There was no evidence for prime surprisal in this part of the dataset either as match had no significant effect (β= -0.04, $\chi2$ (1) = 0.27, p = .61). We then tested the effect following active primes and included bias-match as a fixed effect and by-subject and by-item random slopes. The main effect of prime bias match was not significant (β= -0.05, $\chi2$ (1) = 0.15, p = .71) not providing any evidence for prime surprisal effects in this part of the dataset either. In sum, the transitive prime surprisal study as replicated previous findings of structural priming involving the transitive structure (e.g. Bock, Dell, Chang & Onishi, 2007) but it

[2] We also conducted this analyses with continuous verb-bias predictors (based on passive-frequencies based on both the Manchester corpus and our norming study, see Table 2.2.) where the interaction of prime type and verb-bias would have indicated a prime surprisal effect. None of these analyses changed the pattern of results, neither the main effect of verb-bias, nor its interaction was significant in any of these analyses (all ps > .48).

provided no evidence for prime surprisal.

2.3.3 Assessing alternative explanations for the lack of prime surprisal in the passive dataset

The goal of this study was to assess whether the participants are sensitive to verb bias match in sentences involving transitive (as well as dative) structures and whether this sensitivity is reflected in their structure choices in the target sentences. We found no evidence of prime surprisal in our transitive structures. However, there might be alternative explanations for this lack of effect. Thus, we carried out three exploratory analyses on the transitive dataset to investigate alternative explanations. Our analyses targeted the prime and target sentences and our verb-bias frequency estimates to assess whether they were appropriate in order to yield prime surprisal effects.

First, we tested the possibility that the prime sentences we used were too unusual to be effective primes. As some of the verbs included in the transitive study appeared very infrequently in passive sentences in corpus data (see Table 2.4.), it is possible that participants would perceive passive sentences involving these verbs as ungrammatical, and this would interfere with potential prime surprisal effects. To assess this possibility, we asked 10 native English-speaking students from the University of Liverpool to complete a grammatical acceptability questionnaire. The participants rated the prime sentences from our passive study on a scale from 1 (completely unacceptable) to 7 (completely acceptable). Every participant rated 40 sentences overall, half of the participants rated the primes in an active while the other half in a passive form. Before rating the prime sentences, participants rated seven practice sentences which encouraged them to use the whole scale for their ratings. The sentences were presented in a different order for each participant. The participants received vouchers as a compensation for their participation.

The passive prime sentences received an average overall rating of 6.21 (out of 7) with 64 sentences gaining an average rating of 6 or above, 77 sentences earning a rating of 5 or above and 3 sentences gaining the lowest rating of 4.8. These results suggest that none of the prime sentences were considered ungrammatical and it is not likely that the unusual nature of the primes led to the lack of prime surprisal effects.

Second, we assessed the adequacy of the verb-structure frequencies we used to set up our studies. Similarly to another dative prime surprisal study featuring adult participants (Peter et al. 2015), we used the Manchester corpus (MacWhinney, 2000), a child-directed–speech corpus, to set up our verb-bias groups. Even though the frequencies based on this corpus led to adult prime surprisal effects with dative structures both in Peter and colleagues' and our dative study, we wanted to exclude the possibility that the potential difference between verb-structure frequencies based on a child-directed-speech corpus and those experienced by a student population might have led to the lack of prime surprisal effects in our passive study.

To assess this possibility, we compared the verb-bias frequencies obtained from the Manchester corpus to the frequencies resulting from a norming study featuring student participants that we carried out to set up the event-related potential study described in Chapter 4. In this norming study, 58 students from the Psychology department of the University of California, Davis participated in an online forced choice task where they had to choose between two sentences based on the question: "Which sentence do you find more acceptable?". The study contained 25 transitive sentence pairs in an active and a passive version (e.g. *Bart scared Lisa.* vs. *Lisa was scared by Bart*), and 38 dative sentence pairs in a DOD and a PD version (e.g. *Isabella awarded the nice robin to the chick* Vs. *Isabella awarded the nice chick to the robin*). The study also included 15 catch sentence pairs (e.g. *Isabella danced with* *the goat* vs. *The goat Isabella danced*) where only one of the sentence choices was grammatically correct. We excluded participants who chose more than four incorrect answers in the catch sentence pairs as this indicated that these participants were not reading the sentences carefully before making their choices. We included results from 56 participants in our final dataset. The per verb and per bias-group verb-structure frequencies for the verbs included in our prime surprisal studies based both on the Manchester corpus and on the norming study are featured in Table 2.1. (dative study) and Table 2.2. (transitive study).

Table 2.1.

Verb bias frequencies for the verbs featured in the dative study based on the Manchester corpus and the forced choice norming study.

	Verb	**% DOD in Manchester corpus (per verb)**	**% DOD in Manchester corpus (per category)**	**% DOD in forced choice task (per verb)**	**% DOD in forced choice task (per category)**
PD biased verbs (primes)	Take	15.2	28.44	20.4	28.44
	Sell	23.7		24.1	
	Bring	26.7		40.7	
	Leave	32.7		NA	
	Send	43.9		31.9	
Equibalanced verbs (targets)	Throw	49	34.32	30.9	34.32
	Feed	52.4		33.8	
	Drop	53.3		32.4	
	Flick	55		30.1	
	Slide	56.1		44.4	
DOD biased verbs (primes)	Hand	63.2	80.92	36.6	39.62
	Offer	76.6		43.9	
	Award	82.7		44.9	
	Give	88.8		34.3	
	Show	93.3		38.4	

Table 2.2.
Verb bias frequencies for the verbs featured in the passive study based on the Manchester corpus and the forced choice norming study.

	Verb	% passive in Manchester corpus (per verb)	% passive in Manchester corpus (per category)	% passive in forced choice task (per verb)	% passive in forced choice task (per category)
	Fascinate	26.4		49.5	
	Sadden	16.7		42.6	
More passive biased verbs (primes)	Disgust	30.8	19.78	36.6	39.07
	Dazzle	10		35.3	
	Aggravate	15		31.4	
	Infuriate	4.7		37.9	
	Disturb	6		34.3	
Equibalanced verbs (targets)	Irritate	6.7	5.66	29.4	29.74
	Comfort	5.4		22.1	
	Anger	5.5		25.0	
	Impress	0.07		29.4	
	Shock	1		26.0	
Less passive biased verbs (primes)	Bother	0.09	0.446	25.5	24.71
	Surprise	0.07		21.6	
	Scare	1		21.1	

The verb-bias estimates based on the norming study show a similar pattern to those based on the Manchester corpus. In particular, the sentences including verbs from our more passive biased group were chosen more often as “more adequate” as passive sentences than those featuring verbs featured in our less passive biased group. A similar pattern can be observed in the dative sentences. These results suggest that it is unlikely that the lack of prime surprisal effects are due to that the frequencies determined using the child-directed-speech based Manchester corpus did not map to a student population.

Third and finally, we assessed whether the target verbs featured in the transitive study elicited a high proportion of target responses (coding criteria for the target sentences is discussed in the 2.2.6. Coding section) and therefore had the potential to demonstrate priming and prime surprisal effects. When assessing the proportion of target continuations after each target verb we found that four out of five target verbs led to a high proportion of target continuations. The average target competitions for verbs *comfort*, *disturb*, *infuriate* and *irritate* was 93.7% with the lowest completion rate of 85.4% for the verb *infuriate*. However, the target verb, *anger* led to a very low proportion (31.9%) of target continuations compared to the other target verbs. As this difference suggests that *anger* may not have been an appropriate target verb to demonstrate prime surprisal effects, we wanted to exclude the possibility that the target responses that did include this verb skewed the results and masked a potential prime surprisal effect present with the more appropriate target verbs. So we carried out additional analyses on the transitive dataset excluding responses featuring the target verb *anger*. The first model included (a) prime structure (b) prime-bias match as main effects and by-item random slopes for prime and by-participant random slopes. We found similar response patterns in this dataset to that containing the *anger* as a target verb: we found a significant structural priming effect (β= -0.64, $\chi2$ (1) = 92.63, $p < .00001$) but found no evidence of primes surprisal as neither match (β= -0.01, $\chi2$ (1) = 0.003, p =.95) nor its interaction with prime (β = 0.01, $\chi2$ (1) = 0.05, p = .82) had a significant effect. We also carried out separate analyses on the target sentences following active and passive primes. Both models featured bias-match as a fixed effect and by-subject and by-item random slopes. We found no prime surprisal effects after neither the active (β = -0.01, $\chi2$ (1) = 0.11, p = .91) nor the passive (β = -0.03, $\chi2$ (1) = 0.14, p = .71) primes. As most target verbs elicited a very high proportion of target sentence competitions and the result patterns did not change when excluding the target verb that led to a lower proportion of target responses, it is unlikely that the lack of prime surprisal in the passive study is due to inappropriate target verbs.

2.4. Summary

In summary, we found numerical structural priming effects in the dative and significant structural priming effects in the transitive dataset and also found a non-significant, but numerically large (4.30%) prime surprisal effect in the DOD-priming part of the dative dataset. However, we found no evidence for prime surprisal effects in the study featuring transitive structures and according to our additional analyses the lack of these effects was not due to inadequate prime or target verb selection or verb-structure bias measurements.

2.5. Discussion

We carried out two structural priming experiments in order to gain more information about linguistic predictions by investigating whether prime surprisal effects appear with transitive as well as datives structures. Studies featuring dative structures have demonstrated both structural priming (e.g. Bock, 1986) and prime surprisal effects (e.g. Jaeger & Snider, 2013; Peter et al., 2015). While priming studies featuring transitive structures also demonstrated structural priming effects (e.g. Bock, Dell, Chang & Onishi, 2007), it has not yet been thoroughly investigated whether these structures are also sensitive to prime surprisal.

As expected, the results in the dative dataset were similar to those in previous studies: participants were numerically more likely to produce DOD targets after DOD rather, than PD primes and this priming effect was larger after mis-matching than compared to matching primes. However, the prime surprisal effect was only present in the target sentences following DOD primes, not in those following PD primes, and neither the priming nor the prime surprisal effects reached significance on the .05 level, though they were numerically as large as, or larger than, the structural priming effects in the literature (which tend to be between 3.00 and 4.00%; see Mahowald et al, 2016 for a meta-analysis and see also Peter et.al. 2015 for a dative prime surprisal study similar to ours). As the current prime surprisal effects were numerially comparable to similar effects in the literature and given that we did not conduct a power analysis to determine our sample sizes, we hypothesize that our effects did not reach significance due to our dataset being underpowered.

Another point where our dative results differ from those in previous studies is that we found no prime surprisal effects after PD primes. The literature is not consistent on whether prime surprisal appears after all dative prime sentences or whether only DOD or only PD primes lead to prime surprisal. While Peter et al. (2015) found an overall prime surprisal effect in their adult dataset, Jaeger and Snider (2008) only detected prime surprisal effects after prime sentences featuring PDs, the less frequent prime structure. In our dataset, the DOD was produced less frequently than the PD in our study: overall, 79.89% of the dative target responses were PDs and only 20.91% were DODs. Thus we too found stronger (numerical) prime surprisal effects after the least frequent primes.

Turning to the transitive structure, there was no sign of prime surprisal effects in the transitive dataset. One reason why prime surprisal may be more likely in dative, than transitive structures, is the online processing differences between dative and transitive structures. A critical difference is the location structure decision point (where it becomes clear to the listener which alternative dative or transitive structure are they hearing). In the dative sentences, participants encounter the verb (that carries the information about which structure to predict) before the structure decision point, whereas in transitive sentences the structure is identified before the verb appears, see

Figure 2.4. Thus, considering the incremental nature of online processing, we might argue that continuously predicting one word ahead of the time during the unfolding of a transitive sentence would not lead to either the active or passive structure being surprising based on the verb's bias (as the structure has already been identified by the time the verb appears).

Verb, **Structure decision point**	**Dative sentences**	**Transitive sentences**
Agent/patient is known (picture is present)	Lisa **gave Bart** a ball. Lisa **gave a** ball to Bart.	**Bart scared** Lisa. **Lisa was** scared by Bart.
Agent/patient is unknown (picture is not present)	Lisa **gave Bart** a ball. Lisa **gave a** ball to Bart	Bart **scared** Lisa. Lisa **was scared** by Bart.

Figure 2.4. - Location of the verb (in green) and the structure decision point (in purple) in dative and transitive sentences depending on whether the agent and patient is known before the sentence starts.

Note that in our paradigm, however, the situation is slightly different; we made information about the identity of the verb available from the start, as the verb was written over the pictures the experimenter was describing. So, in our paradigm, sentences that mis-match the prime verb biases should still be more surprising (and potentially lead to larger priming) than matching ones. Thus, the fact that we found no prime surprisal effect in the transitive sentences, suggests that the participants did not consider the verb bias information early on and did not use it to predict the likely structure of the upcoming prime sentence.

It is worth noting, however, that under specific circumstances, transitive sentences could be sensitive to verb-based prime surprisal even if the verb bias information was not available from early on. For instance, if participants encounter a passive structure (e.g. *Dora was…*) and then predict a verb, this verb prediction is more likely to be a passive-biased verb (as passive structures are more often followed by passive-biased verbs). This prediction would then lead to the occurrence of an error-signal if an active biased verb appears instead of a passive biased one. This way mis-matching sentences would be more surprising than matching ones, but this surprisal would be induced by the verb, and not by the structure. Note, however, that this difference could be problematic for the Dual-path model, because if the identity of the verb led to the creation of an error-signal (and not the identity of the structure) this error-signal would be used to strengthen the weights behind the verb (and not the structure) and should not lead to an increased likelihood of structure repetition, i.e. prime surprisal. Furthermore, this route to surprisal would only work under very specific circumstances. For instance, active sentences with passive-biased verbs would only be more surprising if the patient of the sentence was known (otherwise the structure decision point is the verb itself and there would be no space to make verb predictions, see bottom right of Figure 2.4.). Additionally, for prime surprisal with passive sentences it is also necessary that that the verb itself is unknown before the structure decision point (unlike in our study) otherwise the verb can no longer induce surprisal when it appears later on in the sentence. Further work is needed to investigate these possibilities.

However, in the current work there was another crucial difference between the transitive and the dative study that may have influenced the appearance or prime surprisal. Both in previous prime surprisal studies and in our dative study, the two contrasted conditions typically involved sentences that were either DOD- or PD-biased, meaning that in one condition, the first prediction would typically be a DOD structure (and a PD continuation would be surprising) and in the other condition the first prediction would likely to be a PD structure (and a DOD continuation would be surprising). Therefore in these studies the appearance of prime surprisal did not depend on whether participants made any further predictions or not. However, in our transitive study, the two conditions were only differentiated by the relative likelihood of secondary predictions as the two conditions in this study are created by including verbs that are more or less passive biased. In this study alternative graded predictions were necessary for prime surprisal as we included verbs in both conditions for which the first prediction is likely to be the active structure, and the secondary (less likely) prediction would be a passive structure. Consequentially we would only expect participants react to these sentences differently, if they not only make more than one alternative prediction at a time, but they also take into account how likely each of these further predictions are. Again, as we detected no prime surprisal with transitive structures, our study does not provide evidence for the possibility that participants made secondary predictions and taken their likelihood into account for their subsequent sentence production. Importantly though, as there were two crucial differences between the dative and transitive studies that could have either or both have led to a lack of prime surprisal in passives, we cannot determine the exact factors that influenced our results.

To our knowledge, only one other, currently unpublished, study examined prime surprisal effects in transitive sentences. Jaeger and Snider (2008) found larger priming effects after mis-matching (passive sentence, active biased verb) as opposed to matching (passive sentence, passive biased verb) prime sentences in the Treebank Switchboard corpus (Marcus, Santorini, Marcinkiewicz & Taylor, 1999), and this difference only occurred in sentences after passive (but not active) primes. Unlike our study, Jaeger and Snider's work suggest that transitive sentences are sensitive to prime surprisal. This difference in results may be due to one of the differences between the studies. First, Jaeger and Snider's prime verbs were either overall passive- or overall active-biased (based on their verb-structure frequency in the

Treebank corpus), while all our prime verbs were active-biased and only differed in the extent of their active bias (based on whether they appear more often with passives than the other group in the Manchester corpus). Second, participants did not know either the identity of the verb or the identity of the agent and the patient of the sentence before the sentence started (unlike in our study where these were disambiguated by pictures). It is thus possible that in Jaeger and Snider's study the active-biased verb (and not the passive biased structure) led to a prime surprisal, although, as discussed above, this route to prime surprisal does not explain structure repetition according to the Dual-path model. It is also worth noting that Jaeger and Snider mention that their results should be regarded as preliminary and would be followed up in further work (any results of which have not yet been published, to our knowledge). As such, these results cannot be regarded as conclusive in themselves.

2.6. Summary and conclusions

We carried out two prime surprisal studies, one with dative and one with transitive structures in order gain more information about linguistic predictions and their role in error-based learning mechanisms. Our study replicated structural priming effects with both the dative and transitive structures although the differences were only numerical, not significant, for the dative structures. Crucially, we found no evidence of prime surprisal effects in the transitive dataset. As we replicated established structural priming effects with the transitive, it is unlikely that the lack of transitive prime surprisal is due to the shortcomings of the paradigm used. To exclude further non prediction-related explanations, we carried out three additional analyses that confirmed that both the prime and target sentences and the verb-bias measures used here were appropriate to demonstrate transitive prime surprisal effects. In light of the above, we conclude that this study does not provide evidence for prime surprisal with transitive structures in adults.

CHAPTER THREE: DO CHILDREN LEARN FROM THEIR MISTAKES? EVALUATING ERROR-BASED THEORIES OF LANGUAGE ACQUISITION

3.1. Preliminary pilot study

The goal of the main study reported in Chapter 3 below (see section 3.2) was to assess whether less predictable input leads to more long-term as well as more short-term language change. However, we first conducted a pilot study with child participants in order to determine whether the paradigm we were planning to use was suitable for assessing the immediate effects of processing surprising and predictable dative sentences. We also wanted to establish which verbs yielded stable prime surprisal effects in order to create appropriate stimuli.

This pilot study is reported here (section 3.1). The basis of this study was Peter and colleagues' (2015) study that demonstrated immediate prime surprisal effects in the age group also featured in the current study: 5 and 6 year old children. The design, sentence stimuli and procedures were identical to the adult dative prime surprisal study discussed in Chapter 2.

3.1.1. Method

3.1.1.1. Participants

Forty-eight 5-6 year old monolingual English speaking children participated in this study. Five participants were excluded as they produced 'other' responses for more than half of the target trials.

3.1.1.2. Design

The study used a 2x2 design, with prime-type (DOD, double object dative or PD, prepositional dative) and prime-bias match (match or mis-match) as within-subject variables. The dependent variable was the choice of dative structure in the target trials. This study is a partial replication of the study featured in Peter, Chang, Blything and Rowland's 2015 paper, but it featured more verbs and fewer conditions, since our primary goals were to target the immediate prime surprisal effect and identify stimuli for further studies.

3.1.1.3. Visual stimuli

The study featured video animations created in Moho 12, which were presented in the E-prime 2.0 software (Psychology Software Tools, Inc., 2000). The study contained 80 videos overall: 20 and 20 for the prime and target sentences respectively, both depicting transfer actions that can be described with dative sentences, and 40 videos depicting non-causal actions for the filler sentences.

The videos included 10 pairs of donor and recipient characters. Half of them were cartoon characters that are familiar to British children with proper noun names: *Tigger and Piglet*, *Dora (the Explorer) and Boots*, *Marge and Homer*, *Lisa and Bart* and *Bob (the Builder) and Wendy*. The other characters were referred to with determiner and noun NPs: *the prince and the princess*, *the king and the queen*, *the student and the teacher*, *the doctor and the nurse* and *the boy and the girl*. Particular donor and recipient characters were always featured together. A further 10 items acted as objects and were referred to with non-definite determiner and noun NPs: *a monkey*, *a cat*, *a fish*, *a butterfly*, *a fox*, *a bird, a mouse*, *a bunny*, *a baby*, and *a bird*. These characters were consistently paired with one pair of characters (e.g. the *bunny* was always featured with *Bob and Wendy*). In order to keep the intonational contour of the prime and target sentences different, we always featured characters with proper noun names in the target sentences if the primes contained determiner and noun NPs and vice versa (see Rowland et al., 2012).

3.1.1.4. Sentence stimuli

This study contained 80 sentences: 20 prime and target sentences each respectively, and 40 filler sentences. The prime sentences appeared half the time as DOD sentences and half the time as PD sentences. In half of the sentences, the verb bias and the sentence structure matched (DOD-biased verb – DOD structure or PD =-biased verb – PD structure) and in the other half they mis-matched (DOD-biased verb – PD structure or PD-biased verb – DOD structure). The target sentences were produced by the participant based on the video stimuli (as either DOD or PD sentences). The target sentences contained equi-balanced verbs (which appear similarly often in both DOD and PD structures). The filler animations contained non-causal actions that can be described with intransitive sentences. Semi-randomised stimuli lists were created in which the prime and target sentences always followed each other and between each prime-and target pair there were two filler sentences.

The prime and target verbs featured in both studies were selected based on their verb-structure frequencies in the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001, available on CHILDES, MacWhinney, 2000) as this was used as a basis of Peter and colleagues' (2015) prime surprisal study. The Manchester corpus consists of transcripts of audio recordings from a longitudinal study of 12 monolingual English-speaking children between the ages of approximately 2 and 3 years who grew up in major English cities. The verbs are described below with their DOD frequencies (% of datives with this verb that were DODs) based on the Manchester corpus. The first set of verbs was used in the prime sentences. This set contained 5 PD- and 5-DOD biased verbs: *take* (15.2%), *sell* (23.7%), *bring* (26.7%), *leave* (32.7%) and *send* (43.9%) (PD-biased verbs) and *hand* (63.2%), *offer* (76.6%), *award* (82.7%), *give* (88.8%) and *show* (93.3%) (DOD-biased verbs). The second set of verbs was featured in the target sentences. These were equibalanced verbs (that appear with similar frequency in DOD and PD sentences): *throw* (49%), *feed* (52.4%), *drop* (53.3%), *flick* (55%) and *slide* (56.1%).

3.1.1.5. Procedure

The study used the bingo paradigm adapted from Peter et al. (2015). The study was conducted in the form of a bingo game in which the experimenter and the child took turns to describe cartoon animations on a laptop computer. The experimenter and the participant sat in front of the computer side by side. The experimenter introduced all the characters involved in the tasks by showing the participants bingo cards featuring the characters, then described the first cartoon (the prime sentence) and asked the participant to repeat this sentence. The participant was then asked to produce a target sentence by describing a cartoon animation on the other side of the screen. To ensure that the participant's response contained the target verb, a stem-completion technique was used (e.g. the experimented prompted "The king sent..."). After each sentence a smiley or frowny face appeared to signal whether a bingo card was available. If it was available, the child or the experimenter added it to their bingo grid. The first person to fill the bingo grid with bingo cards was the winner of the game and the experiment was designed so that the participant always won. Before starting the study, a practice session was carried out to ensure that the participants understood the task.

3.1.1.6. Coding

The experiment was audiotaped, allowing the transcription and coding of the utterances off-line by the experimenter. A target response was considered a DOD if it contained the correct target verb followed by two noun phrases, and a PD if it contained the correct target verb followed by a noun phrase and a prepositional phrase headed by *to*. Responses were coded as 'other' if (a) the participant failed to repeat the prime correctly (even after help), (b) if the participant produced a non-target verb, or (c) if the target sentence could not be classified as a DOD or PD response based on the above criteria (e.g. target responses containing a preposition other than *to* or incomplete datives such as *the king gave the bunny*).

3.1.2. Statistics and data analyses

The data were analysed by fitting maximal logistic mixed effect models (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). We carried out model comparisons in order to obtain likelihood ratios and p-values. The model comparisons were carried out on maximal models where the random effect structure was only simplified if the model did not converge (Barr, Levy, Sheeper & Tily, 2013). We did not remove any main effects due to non-significance. All the models were calculated using the *lme4* package in R (R Core Team, 2012). The dependent measure was the production of DOD structures (DOD=1, PD=0) in all models. All factors were effect/sum coded (Wendorf, 2004).

We assessed whether we replicated the immediate priming and prime surprisal effects found by Peter and colleagues (2015). Structural priming is demonstrated if there is a greater proportion of DOD responses after DOD rather than PD primes, and immediate prime surprisal effects are demonstrated if there is a significant interaction between prime-structure and prime-bias match, showing that priming effects are larger if the prime verb's bias does not match the prime structure.

The full model included prime-structure and prime-bias match as fixed effects and by-item random intercepts with no random slopes. The main effect of prime structure was significant showing that participants were more likely to produce DODs after DOD rather than PD primes (β= 0.319, $\chi 2$ (1) = 7.77, p < .006). Despite the numerically stronger priming effects after surprising as opposed to predictable prime sentences, the interaction of prime and prime bias match did not reach significance (β= 0.13, $\chi 2$ (1) = 1.22, p = .27) providing no evidence for a significant prime surprisal effect.

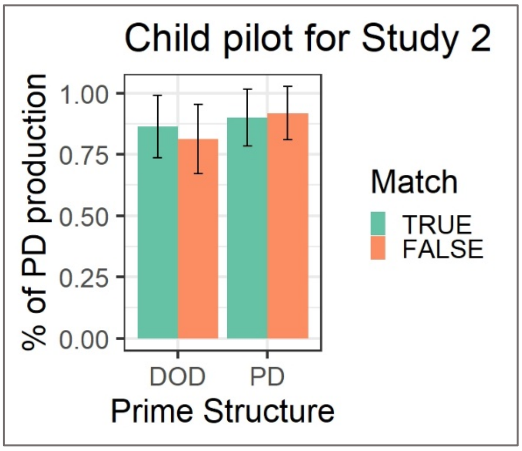

Figure 3.1. – Percentage of PD responses after matching and mis-matching PD and DOD primes. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

Previous prime surprisal studies (e.g. Jaeger & Snider, 2008) found that the prime surprisal effects only occurred after prime sentences featuring the less frequent prime structure, thus we also carried out two sets of analyses on the target structures following the DOD or PD primes separately. The first of these analyses only included the target sentences following DOD primes. This model included prime bias match as a fixed effect and by-participant random slopes. The results revealed no significant prime surprisal effect, although participants were 4.92% more likely to repeat the DOD structure after a mis-matching than after a matching prime (β= 0.19,

$\chi 2$ (1) = 1.65, p = .19). We carried out the second set of analyses on the target sentences following PD primes. This model included prime bias match as a fixed effect and by-participant random slopes. The results revealed no effect of bias match (β= -0.08, $\chi 2$ (1) = 0.19, p = .65) providing no evidence for a prime surprisal effect after PD primes.

3.1.3. Discussion

The goal of the current study was to assess whether the paradigm and stimuli we are planning to use for the study presented in section 3.2. below has the potential to show immediate prime surprisal effects and thus serve as a good basis for extending our examination to longer-term effects as well. Our second goal was to observe how the prime and target verbs behave in this study and use this information to select appropriate stimuli.

The current study showed similar results to our prime surprisal study with adult participants described in Chapter 2. We found a significant structural priming effect but no significant prime surprisal effect in either the full dataset or the subsets involving either only DOD or only PD primes. Despite the lack of significant effects the priming effects were numerically larger in the surprising as opposed to predictable sentences. This effect was driven by the difference in the target sentences following DOD primes where participants were 4.92% more likely to produce DODs after surprising as opposed to predictable primes. There was no such difference in the target sentences after PD primes (where participants were also more likely to produce PDs after surprising PD primes but this difference was only 1.03%).

It is possible that the current study did either not have the appropriate statistical power to detect prime surprisal or that not all verbs featured were suitable to lead to these effects. While Peter and colleagues included 63 5-6 year participants in their study our analyses only featured 42 participants. Furthermore, as one of the goals of this study was to test a wider array of verbs to find the most appropriate ones for our subsequent study, we also included verbs that may be less effective in leading to prime surprisal effects. Thus, the prime surprisal results in our current study may have been weakened by these verbs.

We addressed the above issues for our main study below in two ways. First, we carried out a power calculation based the results in Peter et al (2015) in order to determine the necessary sample size for the main study. Second, we used the information gained about the verbs featured in this study to choose the most appropriate verbs. In order to achieve the second goal we computed how often each verb elicited DODs as featured in DOD or PD primes or as target verbs (see Table 3.1). We used this information when setting up the main study reported in section 3.2. below.

We planned to use the replication of the immediate prime surprisal effects as a contrast for the potential longer terms effects. We used the results of our norming study to ensure the highest probability of the replication. We hoped that replicating an established effect (immediate prime surprisal) while targeting a novel effect (longer-term error-based learning) will help us make more meaningful conclusions about a potential null result for the new effect. We discuss the implication of potential alternative response patterns (involving null results) in the second part of the current chapter.

Table 3.1.

Verb bias frequencies for the verbs featured in the dative study based on the Manchester corpus and the forced choice norming study.

	Verb	**% DOD in Manchester corpus**	**% DOD in target sentences after DOD prime sentences including this verb**	**% DOD in target sentences after PD prime sentences including this verb**	**% DOD in target sentences including this verb**
PD biased verbs	Take	15.2	25.71	10.52	-
	Sell	23.7	16.13	11.11	-
	Bring	26.7	12.5	6.06	-
	Leave	32.7	15.78	11.42	-
	Send	43.9	22.22	11.11	-
Equibalanced verbs	Throw	49	-	-	25.47
	Feed	52.4	-	-	52.89
	Drop	53.3	-	-	16.12
	Flick	55	-	-	58.82
	Slide	56.1	-	-	5.96
DOD biased verbs	Hand	63.2	16.12	13.51	-
	Offer	76.6	9.09	2.94	-
	Award	82.7	7.14	2.77	-
	Give	88.8	10.52	13.51	-
	Show	93.3	22.5	7.69	-

3.2. Main study

3.2.1. Introduction

Prediction, the ability to anticipate other people's upcoming words or actions, plays a key role in a wide range of different human behaviours and activities, from making music (Novembre & Keller, 2011) to playing volleyball (Urgesi, Savonitto, Fabbro & Aglioti, 2012). Prediction plays such a central role in some theories of cognition that human brains have been described as "prediction machines" (Clark, 2013, p. 81). Prediction is particularly important in human communication. It has been suggested, for instance, that prediction can contribute to smooth turn-taking in conversation, not just because it enables us to anticipate when our partner will stop speaking and we can begin speaking ourselves, but also because, by successfully predicting upcoming words, we can give ourselves time to prepare an appropriate response (Magyari, Bastiaansen, de Ruiter & Levinson, 2014). Although some scholars question how central prediction's role in human communication really is (Huettig & Mani, 2016; Hoedemaker & Meyer, 2018), other theories go even further and claim that prediction is a key mechanism in language processing itself (Dell & Chang, 2014; Pickering & Garrod, 2013).

While the role of prediction in adult language use is well documented, there is also the further possibility that prediction is not just vital for using language, but also for acquiring it in childhood. This is the basis of error-based theories of language acquisition. Error-based theories (which can explain learning patterns outside of language as well; Stahl & Feigenson, 2015, O'Reilly, Wyatte & Rohrlich, 2017) suggest that children, like adults, continuously predict upcoming words in conversation, and use these predictions to build up their competence in their first language by comparing what they predicted to the actual input received (Chang, Dell & Bock, 2006; Ramscar, Dye, & McCauley, 2013). One such model, the frequency-based, connectionist Dual-Path model (Chang et al, 2006), uses an error-based learning mechanism (Rumelhart, Hinton, &Williams, 1986) to model the acquisition of syntax, the developmental phenomenon that is the focus of the current study. In this model, if there is a discrepancy between the predicted and actual syntactic structure, an error signal is generated, which is then used to adjust the weights that support syntactic knowledge. These weight changes accumulate over time and allow children's syntactic knowledge to gradually approximate the adult state (note, however, that this is not a stage-based theory; the process also results in representational change in adults, but on a smaller scale since adults' representations are less malleable).

There are several reasons why error-based theories of language acquisition have gained wide support. First, they provide an interactive model that treats language acquisition as the outcome of processing. According to error-based theories, children (and in fact adults) constantly predict words and evaluate predictions while processing language. Every time they make an incorrect prediction, linguistic representations change, which, in children, moves them a step closer to the adult state. This means that error-based theories allow for the possibility that limitations in processing might influence acquisition. Second, rather than simply seeking to define children's state of knowledge at different developmental stages, these models explain how children move from one knowledge state to another. For instance the Dual-Path model not only describes the error-based learning mechanism (that adjusts weights supporting linguistic knowledge in response to error signals), but also demonstrates how this mechanism leads to changes in performance over development (from being able to identify agent and patient roles in intermodal preferential looking experiments at an early age, to producing correct sentences with novel verbs later on). Error-based learning theories thus provide a specific learning mechanism that can be tested experimentally. Third, models implementing error-based learning mechanisms are supported by experimental data and provide explanations for developmental phenomena that are challenging for earlier language acquisition theories. For instance, an error-based noun- acquisition model proposed by Ramscar and colleagues (2013) explains how overgeneralised forms (like “mouses”) disappear from children’s productions in the absence of explicit correction. When children predict the overgeneralized “mouses” form but hear “mice” instead, the associations between the plural of “mouse” and “mouses” weaken due to the error signal resulting from the incorrect prediction, while associations with “mice” are strengthened. Over time, children start producing and predicting the more strongly associated “mice” form instead of “mouses”.

Despite widespread enthusiasm for theories that embrace the role of prediction as a learning mechanism, there remains a major problem. There is, to date, only limited evidence that children actually do generate linguistic predictions, and what evidence there is does not show that these predictions are used for learning. The most promising aspect of error-based theories – that they propose a viable and intuitive language learning mechanism – has therefore yet to be systematically tested. The goal of the present study is to examine the role of prediction in language acquisition by assessing whether less predictable (more surprising) input leads to more lasting change than more predictable input. Below we review the current state of the literature, particularly previous developmental studies on prediction, before discussing the aims of the current project in more detail.

Language acquisition plays a central role in developmental research on prediction, and several experimental studies assess the relationship between prediction and learning. Some studies concentrate on the relationship between predictive abilities and certain aspects of language proficiency (e.g. Mani & Huettig, 2012, Mani, Daum & Huettig 2016; Nation, Marshall & Altmann, 2003, Borovsky & Creel, 2014). For example, Mani and Huettig (2012) found that toddlers' prediction skills (measured using the looking while listening paradigm) significantly correlated with their productive, but not their receptive vocabulary. Other studies have assessed the nature of children's linguistic predictions in order to examine whether they could form the basis of learning (e.g. Borovsky, Elman & Fernald, 2012; Gambi, Pickering & Rabagliati, 2016; Lew-Williams & Fernald, 2007; Lukyaneko & Fisher, 2016). Gambi and colleagues (2016) found that children can use semantic associations as a basis for their predictions (Borovsky et al, 2012) and combine them with predictions based on syntactic knowledge (Huang, Zheng, Meng & Snedeker, 2013), showing that children's predictions could be a viable basis for language acquisition.

Studies targeting prediction in childhood typically use the visual word paradigm, and have been successful in demonstrating that children use anticipatory eye-gazes to visual scenes to predict upcoming words in sentences. However, they do not investigate whether this effect then leads to subsequent learning. They only study whether or not children make predictions; they do not examine if the learning mechanism compares these predictions to actual input or whether the outcome of this comparison leads to subsequent language change. In other words, this paradigm does not address whether predictions form part of an error-based learning mechanism.

There is also another, perhaps more fundamental, problem with using the visual world paradigm to study prediction. This is that so-called predictive looking could, in fact, be a result of a process of integration. In these studies, children listen to sentences where the final word is highly predictable, while their eye-movements on an array of pictures are recorded (Mani & Huettig, 2012; Borovsky et al, 2012; Gambi et al., 2016). Such studies have shown that children as young as 2 years tend to look longer at pictures of objects that would be a more predictable ending for the sentence after hearing the verb, but before hearing the last word (e.g. Mani & Huettig, 212). For example, they are more likely to look at a picture of a cake rather than a picture of a stone after hearing “The boy eats the big …”, that is, before the sentence has been completed. These looks are referred to as anticipatory gazes and are regarded as evidence for prediction. However, according to Rabagliati, Gambi and Pickering (2016), it is possible that these effects are the result of integration and not prediction. If so, children would be looking at the picture of a cake after hearing *eat* because they chose *cake* as the most fitting sentence ending among the given picture alternatives, not because they predicted *cake* themselves. This means that instead of pre-activating upcoming words, children simply incorporate words based on the available visual input (see a similar discussion in the context of EEG research, Kutas & Federmeier, 2011). If so, these studies might not be providing an accurate measure of children’s predictions.

In summary, while some studies have shown a correlation between prediction and learning, and others have shown the potential for prediction to act as a learning mechanism, no studies, to our knowledge, have directly assessed whether predictions lead to lasting changes in underlying linguistic representations – that is, whether they actually contribute to learning in children. In addition, doubts have been expressed in the literature about whether the visual world paradigm really measures prediction or integration.

Our study aims to directly assess both of these issues. We will test whether predictions lead to language learning in childhood using a novel method – prime surprisal (e.g. Jaeger & Snider, 2013; Peter, Chang, Pine, Blything & Rowland, 2015) – to assess whether less predictable linguistic input leads to more lasting language change than more predictable input. This method will not only provide us with information about the immediate and longer-term outcome of correct and incorrect predictions, but will also overcome the problems inherent in using the looking-while-listening paradigm, as it does not involve pictures of more or less predictable sentence endings, and so the responses cannot be guided by visual input.

Prime surprisal studies are based on the priming paradigm (Bock, 1986; Pickering & Branigan, 1998), which is often used to examine syntactic development (e.g. Messenger, Branigan, McLean & Sorace, 2012; Rowland, Chang, Ambridge, Pine & Lieven, 2012). In priming studies, participants are exposed to a prime sentence involving a particular syntactic structure (e.g., active or passive), and then asked to respond to a target stimulus (e.g., a video that they must describe). If participants reuse the previously-processed structure, especially if prime and target sentence share no content, this shows that they have access to the shared (abstract) structural representation underlying the prime and target sentence. This methodology has been particularly useful in demonstrating at what age children develop knowledge of different, abstract syntactic structures. Prime surprisal takes this method a step further by contrasting priming effects in response to predictable and surprising stimuli.

Prime surprisal studies typically feature syntactic structures that can appear in different forms with similar meanings. Dative structures, for instance, appear both as prepositional datives (PD, e.g. *The student gave the report to the teacher*) and double object datives (DOD, e.g. *The student gave the teacher the report*). While DODs appear more often in adult language use overall, every verb has its own specific preferences: for instance, while the verb *give* occurs more often in a DOD structure than in a PD structure, the verb *bring* prefers the PD structure. Children need to acquire these links in order to produce well-formed sentences and avoid incorrect verb- structure pairings (such as '*the student spoke her teacher the answer').

Prime surprisal studies with both child and adult participants have found enhanced priming effects when a structure appeared with a mismatching as opposed to a matching verb (Jaeger & Snider; 2013, Peter et al., 2015). According to the Dual-Path model, these effects result from the error-based prediction mechanism: after hearing a verb, children predict the dative structure that most often follows that particular verb. If they end up hearing a different structure to the one they predicted, the learning mechanism produces an error signal, which they then use to adapt their syntactic knowledge accordingly. In a previous prime surprisal study (Peter et al., 2015), for instance, priming effects were larger when a DOD structure appeared with the verb *bring* (PD-biased) than when it appeared with the verb *give* (DOD-biased), without verb repetition between prime and target sentences. According to the Dual-path model, this occurs because, in the mismatching condition (e.g. DOD with *bring*), participants are likely to make a prediction error. They are likely to predict that the PD-biased verb will be followed by the structure that appears more often with that verb (PD). For example, after hearing "*the boy brings...*" participants are more likely to predict *"... the present to the girl*" (PD) than "*...the girl the present*" (DOD). Since this prediction will turn out to be incorrect, an error signal will be generated, which will, in turn, lead to a change in the weights supporting syntax and to a higher likelihood of the participant reproducing the structure that they have just heard. No such effect occurs in the matching condition: here, when a structure appears with a matching verb (e.g. DOD with *give*), the participants are more likely to successfully predict the upcoming structure, which means that no error signal will be produced. In other words, according to the Dual-Path model, the error signals and weight changes that lead to immediate prime surprisal effects are actually a consequence of the long-term learning that will eventually result in adult-like syntactic preferences.

Although the verb-structure links leading to prime surprisal effects form a key part of syntactic knowledge, they are not fully adult-like at 5 to 6 years of age. According to error-based learning theories, children make predictions from early on, but these early predictions are based on limited linguistic input and therefore are more often incorrect. The older children are, the more adult-like their language becomes, and the more correct predictions they make. At the age of five, children have already accumulated enough knowledge to have verb-structure preferences similar to those of the adults, but since these preferences are based on less linguistic input, they are weaker and more malleable. Children's weaker representations lead to stronger priming effects (Peter et al., 2015) and, according to error-based theories, more learning as well. In contrast, the more developed adult system is less sensitive to the error signals produced by unexpected sentences, resulting in smaller priming and learning effects.

Prime surprisal effects provide promising evidence for prediction in both children and adults, and suggest that incorrect predictions influence subsequent behaviour in the short term. However, the key prediction of this account is that incorrect predictions lead to learning. To test this, we need to demonstrate that prime surprisal leads to lasting cumulative language change as well. To do this, we have developed a new design which combines the prime surprisal method with a paradigm designed to assess whether the original priming effects are cumulative and persistent (see e.g. Kaschak, Kutta & Jones, 2011 for a similar design for adults). These studies typically start with a baseline phase where participants' unbiased rates of the target construction are assessed (e.g. how many DODs and PDs they produce), followed by a test or bias phase where participants are biased towards the production of one of the structures (e.g. are only exposed to PDs or DODs). Finally, in a post-test phase, participants' rates of target construction are re-assessed to see whether they have shifted towards the structure they were biased towards in the previous phase.

Developmental studies using similar designs have shown that children's production frequencies can be shifted towards a less frequent structure by exposure in the bias phase (e.g. Branigan & Messenger, 2016; Kidd, 2012; Savage, Lieven, Theakston & Tomasello, 2006). These results are in line with the predictions of the Dual-Path model, but, due to the set-up of these experiments, they could have originated from sources other than error-based learning. For instance, some studies did not contrast the effects resulting from experience with less-expected structures with the effects resulting from experience with more-expected structures, in which case the post-test shift could be the result of cumulated facilitation from processing a structure multiple times rather than error-based learning (e.g. Kidd, 2012; Savage et al., 2006). Other studies included primes in the post-test phase (as well as the bias phase), meaning that the effects from the bias phase and those of immediate priming are measured on the same target items, making it difficult to tease apart long- and short-term effects (Branigan & Messenger, 2016). The implication is that the strong prediction of the Dual-Path model, that less predictable (i.e. more surprising) linguistic input leads to more lasting language change, still needs to be systematically assessed.

We conducted a four-phase experiment with child and adult participants featuring both predictable and surprising structures in the bias phase, and only including target structures in the baseline and post-test phase. This way, we were able to directly contrast lasting language change resulting from more or less expected structures, and clearly differentiate between immediate and lasting effects of predictability. Furthermore, instead of simply contrasting effects of overall more or less expected structures (e.g. DODs vs. PDs), we also contrasted the effects of the same structure presented in a more or less predictable environment (by consistently presenting PD and DOD structures with either matching or mismatching verbs in a within-participant design, see Table 3.2). This allowed us to get clearer results from the child participants, whose overall dative preferences are inconsistent and not yet adult-like (e.g. Conwell & Demuth, 2007), but who have already been shown to be sensitive to verb-bias effects (Peter et al., 2015). Furthermore, by featuring the same number of PD and DOD structures in both conditions, and only varying how likely it was that participants will correctly predict them, we ensured that the potential differences between results in each condition were due to differences in predictability.

In sum, error-based models that posit prediction as a learning mechanism provide a very promising avenue for understanding the language acquisition process. However, there is limited evidence for the existence of linguistic prediction in childhood, and its contribution to learning has not been systematically examined. To our knowledge, this will be the first study that directly targets the role of prediction in language development by assessing whether unpredictable input leads to more lasting language change than predictable input.

3.2.2. Methods

The goal of this study was to examine the role of error-based learning in acquisition by assessing whether less predictable (more surprising) linguistic input leads to more lasting language change than more predictable input. To achieve this, we used the prime surprisal paradigm in a four-phase experiment, designed to induce error-based learning. The prime surprisal paradigm capitalizes on the fact that some verbs are substantially more likely to appear in one dative sentence structure than another in English, and are thus surprising, despite being grammatical, in the alternative structure. Error-based learning predicts that there will be a bigger change in children's syntactic representations (i.e. learning) after surprising (e.g. PD- biased verb in a DOD structure) than unsurprising (PD-biased verb in a PD structure) primes.

Learning is defined as a change in the underlying syntactic representations and is operationalised as a change in performance from pre- to post- intervention in a production task. More specifically, learning was deemed to have occurred if the children were significantly more likely to use the primed dative structure post-intervention than pre-intervention (i.e. if there was a change in the strength of the children's underlying syntactic representations induced by the priming).

In the first, baseline phase of the study, we assessed participants' baseline rates of dative production (i.e. how many DODs and PDs they produce). Participants described target video animations depicting transitive actions that can be described using dative sentences (e.g. Lisa handing a ball to Bart.), but were free to choose either PD or DOD structures, and the experimenter described filler videos depicting non-causal actions that can be described with intransitive sentences.

The second, priming (or bias) phase was designed to elicit immediate prime surprisal effects, and to bias the participants towards one of the dative structures. Here, participants described target video animations depicting transitive actions in a similar way to the baseline phase, but the experimenter was describing prime animations using either DOD or PD structures. Both structures were consistently paired with either matching or mismatching verbs in the prime sentences (e.g.

DODs only appeared with mis-matching verbs, while PDs only appeared with matching verbs for group A and vice versa for group B). This way, participants in group A were always subjected to PDs in predictable sentences and DODs in surprising sentences.

The third, post-test phase was similar to the baseline phase, but the goal was to reassess participants' rates of dative production. If less predictable input leads to more lasting language change than more predictable input (as suggested by error-based learning theories), participants' production in this phase will shift towards the structure they were exposed to with a mismatching verb in the bias phase (i.e. DODs for participants in group A) compared to the baseline-phase. In order to eliminate the influence of lexically-based long-term priming effects, we used different verbs in the bias and test phases.

While the main focus of this study was abstract error-based learning, the second post-test aimed to assess potential verb-specific lasting priming effects. This phase was similar to the pre- and post-test phases but the target sentences uttered by the participants reused the PD- or DOD-biased verbs that were featured as primes in the bias phase. This way, we would be able to detect a possible change in participants' verb-specific syntactic representations without interfering with the abstract priming effects in the previous phases. If there is verb-specific error-based learning, we expect an enhanced shift towards the dative structure the verb previously appeared with when the structure did not match the verb's bias. For instance, for the PD-biased verb *bring,* we would expect a bigger shift towards the structure for participants for whom it consistently appears with the mismatching DOD structure than for participants for whom it appears with the matching PD structure.

Table 3.2.

General study design showing different trials and verb biases in each phase.

	Group A		Group B	
	Structure	Verb bias	Structure	Verb bias
		Baseline phase		
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
		Test (bias) phase		
Experimenter	**DOD**	**PD-biased**	**DOD**	**DOD-biased**
Participant	Dative	PD-biased	Dative	PD-biased
Experimenter	**PD**	**PD -biased**	**PD**	**DOD-biased**
Participant	Dative	DOD-biased	Dative	DOD-biased
Experimenter	**DOD**	**PD-biased**	**DOD**	**DOD-biased**
Participant	Dative	PD-biased	Dative	PD-biased
Experimenter	**PD**	**PD-biased**	**PD**	**DOD-biased**
Participant	Dative	DOD-biased	Dative	DOD-biased
		Post-test phase		
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
		Second post-test phase		
Experimenter	Filler	NA	Filler	NA
Participant	**Dative**	**PD-biased**	**Dative**	**DOD-biased**
Experimenter	Filler	NA	Filler	NA
Participant	**Dative**	**PD -biased**	**Dative**	**DOD-biased**
Experimenter	Filler	NA	Filler	NA
Participant	**Dative**	**PD-biased**	**Dative**	**DOD-biased**
Experimenter	Filler	NA	Filler	NA
Participant	**Dative**	**PD-biased**	**Dative**	**DOD-biased**

3.2.2.1. Participants

Seventy-two 5- to 6-year-old children (mean age: 72.31 months) and 72 adults, all monolingual English-speakers, participated in the study. The child participants were recruited from schools in the area and an existing participant database, while the adult participants were recruited from the university's student participation pool.

We excluded 10 child and 2 adult participants who produced 'other' responses for more than half of the target trials in the test, post-test or second post-test phases. These participants were replaced in order to obtain 72 sets of data in each age group. Exclusion criteria for the target sentences are discussed in the 2*.2.7. Coding* section below.

These age groups have shown sensitivity to verb-bias manipulations both in the target verb and in the prime verb (prime surprisal) conditions in a priming study involving dative structures (Peter et al., 2015). Children of this age consistently produce both PD and DOD structures (with an average DOD production of approximately 30%) in corpus-based studies (Campbell & Tomasello, 2001) and similar frequencies were observed in priming studies using a similar paradigm to our own (Peter et al., 2015; Rowland et al., 2012). Therefore no floor or ceiling effects were expected to occur in this study.

We determined our sample sizes based on power calculations carried out to allow both of our key comparisons of interest and our manipulation check to be powered adequately. We carried out two sets of power calculations across 1000 iterations on simulated binomial data using mixed effects models, based on those that were used to carry out analyses on our observed data (See *2.2.3. Statistics and data analyses* section). Maximal models were fitted to the simulated data.

If the model failed to converge on 20% of the simulations, it was rejected and simplified before the power analysis was repeated. As our main point of interest in this study was the performance of the child participants, our calculations were based on the effect sizes expected in this group.

Our first power calculation was carried out on our key comparison of interest assessing whether less predictable (more surprising) input leads to more lasting syntactic representation change than predictable input (see power calculation: https://osf.io/9ecjh/ and details of the analyses this calculation is targeting in Section 3.2.1.). As currently there are no data available for our main comparison in the literature, we estimated our simulated effect sizes based on studies targeting contrasts that are in some aspects similar to ours, such as 4 year-olds' post-intervention performance in a study involving the passive structure (Kidd, 2012), an adult intervention study featuring the dative structure (Kaschak, Kutta & Jones, 2011), and a developmental study involving 5 to 6 year olds looking at immediate prime surprisal effects featuring the dative structure (Peter et al., 2015). The effect sizes most relevant to our comparison in the following studies were: 11% post-test shift in a passive intervention study with 4 year olds (Kidd, 2012), an average 7% post-intervention shift in a dative study featuring adults (Kaschak, Kutta & Jones, 2011), and 16% higher priming after mismatching primes than matching ones in an immediate prime surprisal study in 5 to 6 year olds (Peter et al., 2015). Based on the above results we expect at least a 10% shift in both bias groups towards the structure participants were biased towards in the bias phase. In order to ensure that the study was adequately powered even if there were smaller than expected effect sizes, we estimated an average 5% shift in both bias groups (showing that participants' production in the post-test phase will shift towards the structure they were exposed to with a mismatching verb in the bias phase). Based on corpus-based studies (Campbell & Tomasello, 2001) and priming studies using similar materials to our own (Peter et al, 2015.; Rowland et al., 2012), we estimated an average 30% baseline DOD-production in the pre-test phase in both bias groups. Our power calculation showed that our key comparison of interest (post-test differences based on bias group captured by the prime-bias variable) had 93% power when featuring 66 participants. We therefore planned to include 72 participants in each age group in order to have equal numbers of participants in the 8 counterbalance groups and to account for 10% potential data loss.

We also carried out a separate power calculation to ensure that our manipulation check (immediate prime surprisal effect in the test-/or bias-phase, see power calculation: https://osf.io/x2ykf/ and details of the analyses this calculation is targeting in Section 3.1.2.) was adequately powered. As this phase was aiming to replicate the effects in Peter et al.'s 2015 study we simulated data based on the response frequencies in the 5- to 6-year-old group. We estimated an average DOD production of 24% and 35% in the matching PD and DOD prime conditions and 19% and 41% in the mismatching PD and DOD prime conditions. Our power analysis targeted the interaction of prime structure and verb bias. Based on these estimates, the power analysis returned 81.3% power when including 66 participants. With the inclusion of an extra 6 participants (to account for 10% potential data loss), this phase of the study was considered to be sufficiently powered.

3.2.2.2. Design

The between-subject variables were age (adults vs. children) and prime bias (DOD-bias and PD-bias), and the within-subject variables were verb-bias match (match or mismatch), prime-type (DOD and PD) and phase (pre-test, bias phase, post-test and second post-test). The dependent variable was the choice of dative structure in the target trials.

3.2.2.3. Predictions

We had four main predications, which will be discussed in more detail in Section 2.2.3 (Statistics and data analyses).

1. Immediate prime surprisal: we expected to replicate the effects found in Peter et al.'s (2015) study and find increased priming if the verb bias and the prime structure did not match in the prime sentence.

2. Learning about abstract structures: we expected that less predictable (more surprising) input will lead to more learning than predictable input. Therefore, participants' production in the post-test should shift towards the structure they were exposed to with mismatching verbs in the bias phase.

3. Verb based learning: due to the larger learning effects resulting from unpredicted input, we expected that participants would be more likely to reuse the structure the target verb previously appeared in if that structure did not match the verb's bias.

4. Stronger effects in the child than in the adult group: due to the weaker and more malleable verb-biases in children compared to adults, we expected that the three above effects (immediate prime surprisal, learning about abstract structures and verb-based learning) would be larger for children than adults.

3.2.2.4. Visual stimuli

The study featured video animations created in Moho 12, which were presented in E- prime 2.0 software (Psychology Software Tools, Inc., 2000). Each participant saw 120 videos: 60 videos depicting transfer actions that can be described with prepositional or double object datives for the prime and target sentences and 60 videos depicting non-causal actions for the filler sentences.

The cartoons included 10 pairs of donor and recipient characters. Half of them were cartoon characters that are familiar to British children with proper noun names: *Tigger and Piglet*, *Dora (the Explorer) and Boots*, *Marge and Homer*, *Lisa and Bart* and *Bob (the Builder) and Wendy*. The other characters were referred to with determiner and noun NPs: *the prince and the princess*, *the king and the queen*, *the student and the teacher*, *the doctor and the nurse* and *the boy and the girl*. Particular donor and recipient characters were always featured together. A further 10 items acted as objects and were referred to with non-definite determiner and noun NPs: *a ball*, *a toy*, *an orange*, *a cake*, *a peach*, *a sandwich, a pencil*, *a book*, *a napkin*, and *an apple*. The objects were consistently paired with one pair of characters (e.g. the *ball* was always featured with *Bob and Wendy*).

In the bias phase, prime videos were always paired with a target video that included different characters from those in the prime. In order to control for the possibility that direction of transfer may influence structure choice, the animations depicted the direction of motion of transfer actions equally often from right-to-left and from left-to-right.

3.2.2.5. Sentence stimuli.

The study contained 120 sentences (including 60 verb stems) per participant: 16 prime and 16 target sentences plus 32 fillers in the bias phase, 10 target and 10 filler sentences in the pre- and post-test phases and 8 target and 8 filler sentences in the second post-test. The prime sentences appeared half the time as DOD sentences and half the time as PD sentences. Both structures were consistently paired with either matching or mismatching verbs in the prime sentences (e.g. PDs only appeared with matching verbs while DODs only appeared with mismatching verbs for participant A and vice versa for participant B). The target sentences were produced by the participant (as either DOD or PD sentences) based on the video stimuli.

For instance, a prime-target trial in the bias phase included a prime sentence such as *The king brought the queen a cat.* (DOD) or *The king brought the cat to the queen.* (PD) and participants completed a sentence stem such as "Lisa dropped…" as a target sentence.

In order to avoid lexically-based long-term priming effects, we used a different set of verbs in the bias phase- and in the pre- and post-test phases. The study involved the following two sets of verbs, featured here with their DOD frequencies in the Manchester corpus (Theakston, Lieven, Pine & Rowland, 2001, available on CHILDES, MacWhinney, 2000) in brackets (for the computation of the dative frequencies see Ambridge, Pine, Rowland, Freudenthal & Chang, 2014). The first set of verbs was used in the pre- and post-test phases. This set contained 3 equi-balanced verbs: *feed* (52%), *slide* (56%), and *throw* (49%), and one PD- and one DOD-biased verb: *bring* (27%) and *give* (89%). The second set of verbs was featured in the test-phase and repeated in the second post-test. This set contained four PD- biased verbs: *leave* (32%), *sell* (24%), *send* (44%) and *take* (15%) and four DOD-biased verbs: *award* (83%), *hand* (63%), *offer* (77%) and *show* (93%).

We selected the above verbs based on the frequency of their dative occurrences in the Manchester corpus. These verbs have yielded immediate prime surprisal effects in other studies featuring similar age-groups to ours (Peter et al., 2015) as well in our pilot study featuring 5- to 6-year-old children. We aimed to select verbs that have strong verb biases for the bias-phase (as prime surprisal is defined as the negative logarithm of the verb bias, Fine, Jaeger, Farmer & Quian, 2013), but our choices were constrained by the limited number of verbs that appeared often in dative structures in the Manchester corpus.

To control for sentence-specific preferences, we created eight counterbalance groups to ensure that 1. if the DOD structure consistently appeared with matching verbs in one counterbalance group, it appeared with mismatching verbs in the other (and vice-versa for the PD structure), and 2. if a verb appeared with a DOD in a counterbalance group, it appeared with a PD in the other and 3. if a target sentence appeared in the pre-test in one counterbalance group, it appeared in the post-test in the other.

Semi-randomised stimulus lists were created in which the prime and target sentences always followed each other in the bias phase and the same verb did not appear twice in immediate succession. In the test- or bias-phase, there was always a pair of filler sentences after every target sentence. In the other phases, filler and target phrases alternated with each other.

3.2.2.6. Procedure

The study used the bingo game paradigm (Peter et al., 2015, Rowland et al., 2012). It takes the form of a bingo game in which experimenter and child take turns to describe cartoon animations or pictures on a laptop computer. The experimenter introduced the characters involved in the tasks by showing the participants cards featuring the characters. The experimenter and the participant sat in front of the computer side by side. The experimenter described the first cartoon and asked the participant to repeat the sentence. The participant was then asked to produce a target sentence by describing a cartoon animation on the other side of the screen. To ensure that participants' responses contain the target verb, a stem-completion technique was used (e.g. *the boy gave...*). Each target sentence was immediately followed by an intransitive filler sentence.

After every two sentences, a smiley or frowny face appeared in order to signal whether a bingo card was available. If it was, the child or the experimenter received the card and added it to their bingo grid. The first person to fill the bingo grid with bingo cards was the winner of the game, and the experiment was designed so that the participant always won.

Before beginning the study, there was a practice session to ensure that the participants understood the task. The practice session included intransitive sentences featuring three characters each (e.g. "The king and the queen were playing with the cat."). In order to encourage the production of full datives in the main study, we asked participants to mention all three characters in their descriptions during the practice session. To further encourage the production of full datives in the study, the first verb featured as a target in both the pre- and post-test phase was a verb that cannot be used as an intransitive.

The bingo paradigm paired with the stem-completion technique has been successfully used to elicit dative sentences in similar age groups and has resulted in low exclusion rates (Peter et al., 2015; Rowland et al., 2012). Furthermore, both the child and adult participants enjoyed participating in our pilot study featuring this paradigm and all participants completed the session.

After children completed the Bingo game we administered the Test for Reception of Grammar to measure their morphosyntactic abilities. The study lasted approximately 40 minutes, including a break, and participants received a sticker after the practice session.

3.2.2.7. Coding

The experiment was audiotaped, allowing the transcription and coding of the utterances off-line by the first author. A target response was considered a DOD if it contained the correct target verb followed by two noun phrases, and a PD if it contained the correct target verb followed by a noun phrase and a prepositional phrase headed by *to*. Responses were coded as 'other' if (a) the participant failed to repeat the prime correctly (even after help), (b) if the participant produced a non-target verb, or (c) if the target sentence could not be classified as a DOD or PD response based on the above criteria (e.g. target responses containing a preposition other than *to* or incomplete datives such as *the king gave the ball*).

3.2.3. Statistics and data analyses

The data was analysed by fitting maximal logistic mixed effect models (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). We carried out model comparisons in order to obtain likelihood ratios and p-values. The model comparisons were carried out on maximal models where the random effect structure was only simplified if the model did not converge (Barr, Levy, Sheeper & Tily, 2013). We did not remove any main effects due to non-significance. All the models were calculated using the *lme4* package in R (R Core Team, 2012). The dependent measure was the production of DOD structures (DOD=1, PD=0) in all models. All factors were effect/sum coded (Wendorf, 2004) excluding age in months and TROG score, which were centered (Neter, Wasserman & Kutner, 1985).

We carried out three sets of analyses on different sections of the response data to (1) assess whether we replicated the prime surprisal effects found in Peter and colleagues' (2015) study, (2) explore whether less predictable (more surprising) linguistic input led to more lasting language change than more predictable input with no repetition of verbs; and (3) explore whether less predictable (more surprising) linguistic input led to more lasting language change than more predictable input for repeated verbs. In the following sections, we describe all analyses involving data from both age groups together, but, in order to explore the group specific patterns in more detail, we also carried out analyses on the data from the two age groups separately. The main effect of age and TROG score in months were added to the models examining data from the child group.

3.2.3.1. Confirming expected effects: Hypothesis 1– Immediate priming effects are increased if the prime structure appeared with a mismatching as opposed to a matching verb (immediate prime surprisal effect)

This analysis served as a manipulation check. To confirm the differences in predictability between the different bias conditions (that were designed to lead to long-term changes in the post-test phase), we first assessed whether we replicated the immediate prime surprisal effects found by Peter and colleagues (2015). Immediate structural priming is demonstrated if there is a greater proportion of DOD responses after DOD than PD primes, and immediate prime surprisal effects are demonstrated if there is a significant interaction between prime-structure and prime-bias match, showing that priming effects are larger if the prime verb's bias does not match the prime structure.

3.2.3.1.1. Both age groups

The first analysis was carried out on the target sentences from the bias phase including data from both adult and child participants. The full model included as fixed effects: prime-structure (DOD or PD), prime-bias match (depending on whether the prime verb's bias matched or mismatched the prime structure) and age group (children or adults), and by-item random intercepts with no random slopes. We found a main effect of age group, indicating that overall children produce fewer DODs than adults (β= 0.91, $\chi2$ (1) = 10.39, $p < .0001$) and a main effect of prime, indicating that participants were more likely to produce DODs after DOD than PD primes (β= 0.11, $\chi2$ (1) = 4.25, $p < .04$). The interaction of prime and age group did not reach significance (β= -0.09, $\chi2$ (1) = 3.22, p = .09), so this dataset provided no evidence for a difference in overall priming effects between the two age groups.

In terms of prime surprisal, there was a numerically larger priming effect after mis-matching (surprising) than matching (predictable) primes, but this interaction of match and prime did not reach significance (β= 0.07, $\chi2$ (1) = 3.01, p = .21). However, the three-way interaction of prime-bias match, prime structure and age group was significant (β= 0.11, $\chi2$ (1) = 4.6, $p < .04$) indicating a larger prime surprisal effect in the adults than in the children. (β= 07, $\chi2$ (1) = 0.19, p =.65). In order to explore the group specific patterns in more detail, we also carried out analyses on the data from the two age groups separately.

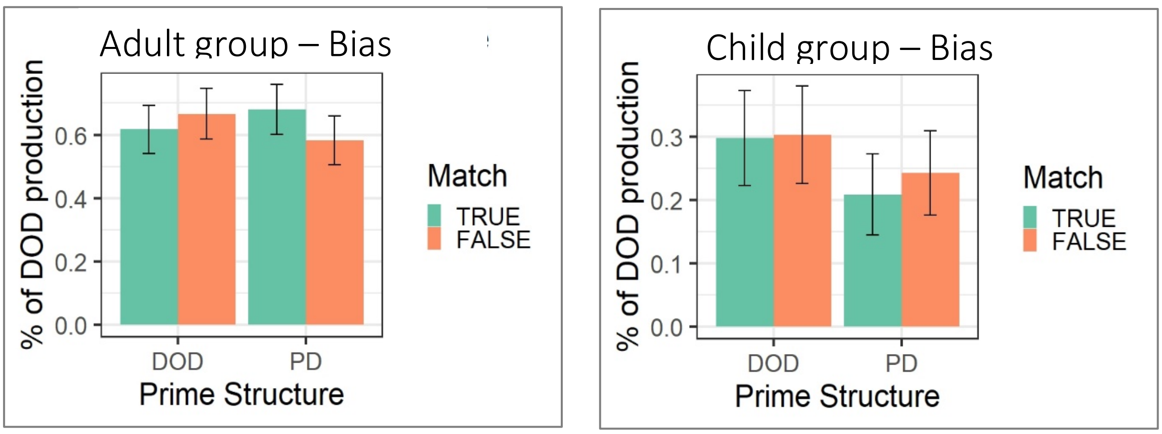

Figure 3.2. – Proportion of DOD production in the bias phase in both age groups. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

3.2.3.1.2. Adult group

The full model included prime-structure and prime-bias match as fixed effects and by-item and by-subject random intercepts with no random slopes. None of the main effects of prime structure (β= 0.032, $\chi2$ (1) = 0.165, p = .68) or prime bias match (β= -0.62, $\chi2$ (1) = 0.604, p = .43) reached significance. Crucially, despite the numerically stronger priming effects after surprising as opposed to predictable prime sentences, the interaction of prime and prime bias match did not reach significance in this analysis (β= -0.28, $\chi2$ (1) = 1.73, p = .185) providing no evidence for a significant prime surprisal effect in the adult group.

3.2.3.1.3. Child group

In the child group we first carried out the analyses modelled in the power calculation discussed in Section 2.1. The full model included prime-structure and prime-bias match as fixed effects and by-item and by-subject random intercepts with no random slopes. The main effect of prime was significant (β= 0.26, $\chi2$ (1) = 7.62, $p < .05$). However, neither prime bias match (β= -0.23, $\chi2$ (1) = 0.45, p = .62) nor its interaction with prime (β= -0.07, $\chi2$ (1) = 0.11, p = .74) reached significance in this analysis.

In order to explore the effect of age and syntactic knowledge, we built an additional model where we included age in months and TROG score in addition to the main effects of match and prime. The model also included by-item random intercepts with no random slopes. In this model, prime (β= 0.21, $\chi2$ (1) = 6.2, $p < .02$), TROG score (β= 0.43, $\chi2$ (1) = 19.52, $p < .0001$) and age (β= -0.288, $\chi2$ (1) = 0.47, $p < .003$) had main effects, demonstrating a significant priming effect and showing that younger children and those with a higher TROG score were more likely to produce DOD structures overall. None of the other main effects or interactions reached significance (all p values > 0.34). Importantly, the interaction of prime and bias match which would have indicated a prime surprisal effect did not reach significance (β= -0.96, $\chi 2$ (1) =0.01, p = .95).

3.2.3.1.4. Summary of bias-phase results

The priming effects in the adult group were larger after surprising as opposed to predictable primes as expected, but neither the priming nor the prime surprisal effect reached significance. In the child group, while there was a significant priming effect, there was no sign of prime surprisal. In fact, children were slightly (but not significantly) less likely to repeat the prime structure when it appeared in an expected as opposed to unexpected sentence structure.

As the main goal of this phase was to induce different rates of error-based learning, we used a partially between-participant design that led to differences compared to other prime surprisal studies. To assess whether these differences contributed to the absence of prime surprisal we carried out three additional analyses to address whether (1) the variable baseline performance, (2) participants only hearing sentences that lead to either DOD or a PD bias or (3) the increasing predictability of the verb-structure pairings influenced the prime surprisal effects. We carried out these analyses in both age groups.

3.2.3.1.5. Exploratory analyses

3.2.3.1.5.1. Exploring the effects of baseline DOD performance

To address the possibility that the baseline DOD-production rates influenced the prime surprisal effects, we carried out an additional analysis that included baseline DOD production (as measured in the pre-test phase) in the models.

3.2.3.1.5.1.1. Adult group

The full model included prime-structure, prime-bias match and baseline DOD production as fixed effects and by-item and by-subject random intercepts with no random slopes. There was a significant main effect of baseline DOD production (β= 2.66, $\chi 2$ (1) = 30.12, $p < .00001$) but neither the effect of match (β= -0.06, $\chi 2$ (1) = 0.6, $p = .43$) nor prime structure (β= 0.03, $\chi 2$ (1) = 0.16, $p = .68$) reached significance, providing no evidence for a priming effect. However, the interaction of match and prime had a significant effect in this analyses (β= 0.275, $\chi 2$ (1) = 4.00, $p < .05$) showing a significant prime surprisal effect in the adult group when baseline DOD production was taken into account.

3.2.3.1.5.1.2. Child group

The full model included prime-structure, prime-bias match and baseline DOD production as fixed effects and by-item and by-subject random intercepts with no random slopes. The main effect of prime structure (β= 0.22, $\chi 2$ (1) = 6.20, $p < .001$) and baseline DOD production (β= 0.766, $\chi 2$ (1) = 87.71, $p < .0001$) were significant but the main effect of match (β= 0.04, $\chi 2$ (1) = 0.36, $p > .52$) was not. Crucially, even though the positive beta value indicates a larger priming effect after surprising as opposed to predictable primes, the interaction of prime and verb bias match did not reach significance (β= 0.93, $\chi 2$ (1) = 1.17, $p > .28$). These results show no significant prime surprisal effect in the child group even when taking baseline DOD performance into account.

3.2.3.1.5.2. Exploring the effects of bias group

In previous fully within-participant prime surprisal studies, participants heard sentences from four conditions, two of which were more likely to lead to DOD bias (surprising DOD, predictable PD) while the other two were more likely to lead to a PD bias (surprising PD, predictable DOD). Thus throughout the study participants overall PD versus DOD-bias was roughly equal. However as the main goal of our bias phase was to bias participants towards more DOD or PD-production they only heard two out of four conditions each. Thus participants' cumulative experience was overall already more DOD- or PD-biased in the bias-phase. This difference compared to previous prime surprisal studies may have interfered with the immediate prime surprisal results. To assess this possibility, we added overall bias group to the previous analyses.

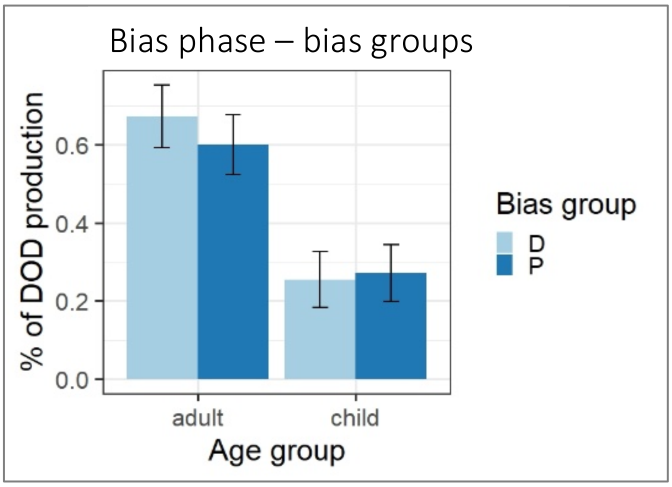

Figure 3.3. – Proportion of DOD production in the bias-phase per overall bias group. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

3.2.3.1.5.2.1. Adult group

The full model included prime-structure and bias group as fixed effects and by-item and by-subject random intercepts with no random slopes. While adult participants were 7.31% more likely to produce DODs in the DOD- as opposed to the PD bias group in the bias-phase, bias group had no significant effect ($\beta = 0.16$, $\chi2$ (1) = 1.25, $p = .28$). This analysis provided no evidence that cumulative experience significantly influenced the participants in the bias phase. None of the other effects reached significance either (all p values > 0.54).

3.2.3.1.5.2.2. Child group

The full model in the child group included prime-structure and bias group as fixed effects and by-item and by-subject random intercepts with no random slopes. As in previous analyses in the child group, we found a significant priming effect (β= 0.19, $\chi2$ (1) = 6.28, $p < .02$). However bias group had no significant effect in this age group either (β= 0.16, $\chi2$ (1) = 1.25, $p = .28$). None of the other effects reached significance (all p values > 0.54).

3.2.3.1.5.3. Exploring the effects of decreasing level of surprisal for repeated verb-structure pairings

To allow us to assess verb-based learning effects in the later stages of the study, we consistently paired each verb with the same dative structure throughout the bias-phase for each participant. This meant that the specific verb-structure pairings become more and more predictable and may not have led to prime surprisal effects later on in the bias-phase. To address this possibility, we carried out an additional analysis that only included the first two occurrences for each verb, where the verb-structure pairing had not yet become predictable.

3.2.3.1.5.3.1. Adult group

The full model included prime-structure and prime-bias match as fixed effects and by-item and by-subject random intercepts with no random slopes. None of the main effects of prime structure (β= 0.12, $\chi2$ (1) = 1.07, $p = .28$) or prime bias match (β= 0.06, $\chi2$ (1) = 0.55, $p = .58$) nor their interaction (β= -0.14, $\chi2$ (1) = 1.37, $p = .42$) reached significance in this analyses.

3.2.3.1.5.3.2. Child group

The full model included prime-structure and prime-bias match as fixed effects and by-item and by-subject random intercepts with no random slopes. The main effect of prime structure remained significant in this analysis (β= 1.53, $\chi2$ (1) = 9.26, p =

.002). However, while children showed 2.32% more priming after surprising as opposed to predictable sentences, the interaction of prime structure and verb bias match did not reach significance in this analysis either (β= 0.1, $\chi2$ (1) = 1.73, p = .48).

3.2.3.2. Key comparison of interest: H2a – participants' production in the first post-test phase shifts towards the structure they were exposed to with mismatching verbs in the bias phase

The second analysis tested the main prediction of error-based learning: that less predictable (more surprising) input leads to more lasting syntactic representational change than predictable input. This analysis assessed whether the post-test scores differ in the two bias-groups, while controlling for the pre-test performance.

3.2.3.2.1. Both age groups

The first analyses was carried out on the target items from the post-test phase and the full model included bias-group (depending on whether participants were biased towards DOD or PD structures in the bias-phase), pre-test score (the proportion of DODs for all dative structures a participant produced in the pre-test phase) and age group (children or adults) as fixed effects, and by-subject and by-item random intercept with no random slopes. There was a main effect of age group (β= 0.43, $\chi2$ (1) = 37.53, p < .02), indicating that children were, overall, less likely to produce DODs than adults. Pre-test score also had a main effect (β= 4.99, $\chi2$ (1) = 67.52, p < .0001) showing that participants with higher baseline DOD production (measured in the pre-test phase) were overall more likely to produce DODs in the post-test phase as well. However, while participants shifted towards the structure they were biased towards in both groups (average 4.25% shift in the adult and average 6% shift in the child group), the main effect of bias group did not reach significance (β= 0.41, $\chi2$ (1)

= 0.28, p = .59). None of the other main effects reached significance either (all ps > 0.53). Despite a numerically larger shift in the child than in the adult group, the interaction of bias-group and age-group was not significant either (β= -0.07, $\chi2$ (1) = 0.19, p =.65). In order to explore the group-specific patterns in more detail, we also carried out analyses on the data from the two age groups separately.

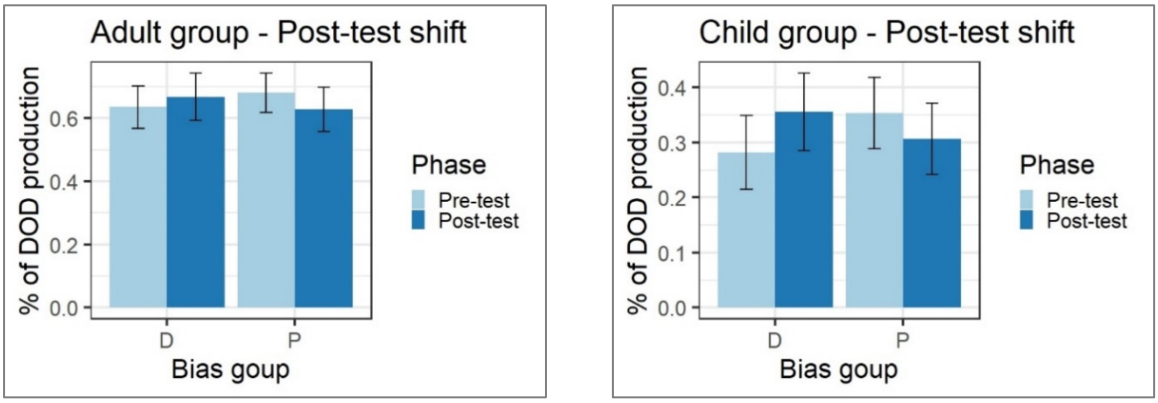

Figure 3.4. – Pre-to post test shift in both age groups. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

3.2.3.2.2. Adult group

The model fitted to the post-test scores from the adult group included bias group and pre-test-score as fixed effects and by-subject and by-item random intercept with no random slopes. Pre-test score had a significant positive effect on post-test scores (β= 4.42, $\chi2$ (1) =30.47, p < .0001). However, the main effect of bias group did not reach significance in the adult dataset (β= 0.23, $\chi2$ (1) = 0.83, p = .24), although there was a numerical difference (on average adults produced 3.22% more DODs in the DOD- and 5.33% more PD in the PD-bias groups compared to their baseline production).

3.2.3.2.3. Child group

The first model fitted to the child group separately was a replication of the analyses modelled in the power calculation described in section 2.1.. This analysis included bias group and pre-test-score as fixed effects and by-subject and by-item random intercept with no random slopes. Pre-test score had a significant positive main effect (β= 1.47, $\chi2$ (1) = 27.16, p < .0001). As we found for adults, children produced 7.33% more DODs in the DOD- and 4.72% more PD in the PD-bias group compared to their baseline production, but this main effect of bias group did not reach significance (β= 0.42, $\chi2$ (1) = 0.35, p = .55).

To assess the effect of age and syntactic knowledge, we built an additional model where we included age in months and TROG score as main effects in addition to bias group and pre-test score and by-subject and by-item random intercept with no random slopes. In this model, only pre-test performance had a significant effect (β= 1.54, $\chi2$ (1) = 29.26, p < .0001). None of the other effects reached significance (all p values > 0.106). Importantly, neither the main effect of bias group (β= 0.39, $\chi2$ (1) = 2.21, p = .14), nor its interaction with TROG score (β= 0.16, $\chi2$ (1) = 0.03, p =.95) or age in months (β= -0.45, $\chi2$ (1) = 2.61, p = .11) was significant.

3.2.3.2.4. Summary of the results of the post-test-phase

In summary, both age groups showed a pre- to post-test shift towards the dative structure they were exposed to in a surprising (as opposed to predictable) sentences in the bias phase, but these results did not reach significance in either the full dataset including both age groups or in either of the age groups separately. The reason for the lack of significant results may be either the smaller than anticipated effect sizes (in the adult group), the larger variance in DOD-production (in the child group) or the ceiling performance some participants showed in the pre-test phase. We will address the last possibility in the next section.

3.2.3.2.5. Exploratory bias phase analyses without participants who showed a

ceiling performance in the pre-test phase

In order to adequately measure the effect of our main manipulation (positive or negative pre- to post-test shift depending on bias group), it is essential that participants do not show ceiling performance in the pre-test phase. For instance if a participant already produces 100% DODs in the pre-test phase, they can only shift towards higher PD (and not higher DOD) production in the post-test phase. While participants on average produced 32.52% DODs in the child group and 62.32% DODs in the adult group, some participants (14 children and 20 adults) produced exclusively PDs or exclusively DODs in the pre-test. Thus, we conducted a separate analysis including only the participants who did not show ceiling performance in the pre-test phase. We replicated the previously discussed analyses on this dataset.

3.2.3.2.5.1. Both age-groups together

The first analysis was carried out on the target items from the post-test phase in both age groups, excluding participants who showed ceiling performance in the pre-test phase. This analysis included 110 participants (52 adults and 58 children). The full model included bias-group, pre-test score and age-group as fixed effects and by-subject and by-item random intercept with no random slopes. We found a main effect of age group (β= 0.29, $\chi 2$ (1) = 63.42, $p > .0001$), pre-test score (β= 3.00, $\chi 2$ (1) = 95.94, $p < .0001$) and crucially a main effect of bias group (β= 0.26, $\chi 2$ (1) = 7.22, $p < .0005$), showing that participants were significantly more likely to produce DODs in the DOD as opposed to the PD bias group. Despite the numerically larger average pre- to post-test shift in the child compared to the adult group (average 4.23% shift in the adult and average 7.91% shift in the child group), the interaction of bias-group and age-group did not reach significance in this analyses (β= -0.06, $\chi 2$ (1) = 0.67, p =.41).

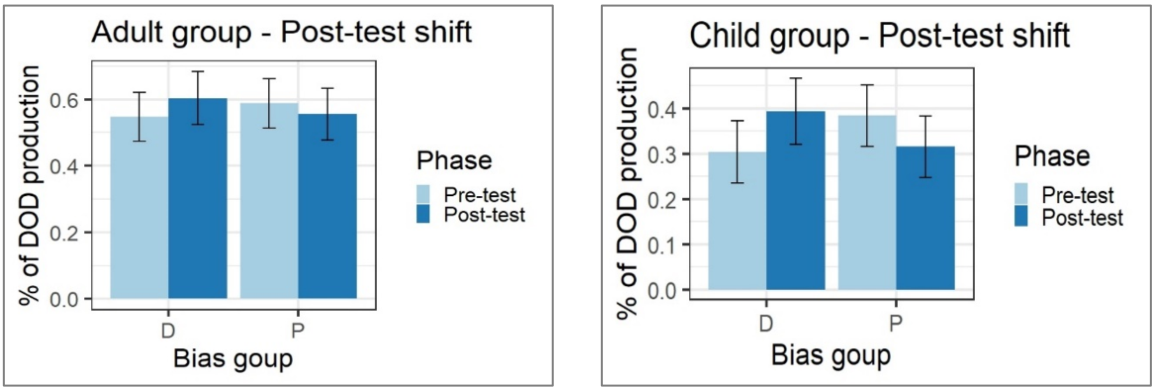

Figure 3.5. – Pre-to post test shift in both age groups without participants with ceiling pre-test values. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

3.2.3.2.5.2. Adult group

This analysis included 52 adult participants. The model included bias group and pre-test-score as fixed effects and by-subject and by-item random intercept with no random slopes. Pre-test score had a main effect (β= 2.98, $\chi2$ (1) =19.24, $p < .0003$) but the main effect of bias group did not reach significance (β= 0.19, $\chi2$ (1) = 0.33, $p = .62$) showing that despite the 5.52% shift in the DOD and 2.91% in the PD bias group, the pre- to post-test shift was not significant in the adult group.

3.2.3.2.5.3. Child group

This analysis included 58 child participants. The first analysis in the child group included bias group and pre-test-score as fixed effects and by-subject and by-item random intercept with no random slopes. Pre-test score had a main effect (β= 1.14, $\chi2$ (1) = 17.13, $p < .0001$). As with the adults, the children produced 8.94% more DODs in the DOD- and 6.86% more PD in the PD-bias group compared to their baseline production but the main effect of bias group did not reach significance (β= 0.49, $\chi2$ (1) = 0.94, $p = .33$).

To explore the contribution of age and syntactic knowledge, we built an additional model where we included age in months and TROG score in addition to pre-test score and bias group. We found a main effect of pre-test score ($\beta= 0.82$, $\chi2$ (1) = 49.13, $p < .0001$). Importantly the main effect of bias group was also significant in this analyses ($\beta= 0.78$, $\chi2$ (1) = 12.87, $p < .004$), showing that children produced significantly more DODs in the DOD than in the PD bias group when age and syntactic knowledge were taken into account. The interaction of bias group and age also reached significance ($\beta= -0.34$, $\chi2$ (1) = 8.32, $p < .04$) indicating that younger children were more likely to be influenced by our manipulation in the bias phase. None of the other effects reached significance (all p values > 0.08).

3.2.3.2.5.4. Summary of the exploratory analyses in the post-test phase

In summary, the full dataset and the subset dataset (that included only participants with no pre-test ceiling effect) yielded a similar pattern of results. However while none of the effects reached significance in the full dataset, in the subset dataset both participants overall, and participants in the child group separately showed significantly different rates of DOD production, in the predicted direction, depending on the bias group they were in. The effect in the child group however only reached significance when syntactic knowledge and age were taken into account.

3.2.3.3. Additional analyses of potential interest: Hypothesis 3 - Verb-based long-term effects of input predictability

The third set of analyses was carried out on the target sentences from the second post-test phase and the goal was to detect verb-specific long-term learning effects. The Dual-Path model predicts both abstract and verb-based learning effects. When participants hear an unexpected structure (given the verb's bias) they use the resulting error signal to increase the weights supporting both the unexpected structure (e.g.

DOD) and the structure's connection with the verb (e.g. the connection between DOD and *sell*). In line with this prediction we expect to see a main effect of prime structure showing that participants are more likely to reuse the structure in which they previously heard the verb. Crucially, a lasting verb-specific prime surprisal effect would be demonstrated by an interaction between prime-structure and verb-bias match showing that participants are more likely to reuse the structure the target verb previously appeared in if that structure did not match the verb's bias.

Unlike the data described in H2, we do not have a baseline for comparing performance in the second post-test for this analysis. Furthermore, the multiple exposure to these verbs in the different phases may influence the prime surprisal effects. Consequently, this analysis provides a secondary test of prediction-based learning. The experiment was primarily designed to look for abstract, structural effects whereas the purpose of this analysis was to examine possible verb- specific effects in the data, which are hypothesized in error-based theories implementing learning via prediction.

3.2.3.3.1. Both age-groups together

The full model included prime-structure (PD or DOD depending on which structure the verb appeared in during the bias phase), verb-bias match (depending on whether the verb featured here appeared in a matching or mis-matching sentence), target verb bias (proportion of DOD occurrence in the Manchester corpus) and age-group (adults or children) as fixed effects and by-subject random intercepts with no random slopes. Prime structure (β= 0.37, $\chi2$ (1) = 13.12, $p < .0001$), age group (β= 1.77, $\chi2$ (1) = 42.79, $p < .0001$) and verb bias match (β= 3.16, $\chi2$ (1) = 28.37, $p < .0001$) had a significant effect, showing that participants in the adult group were more likely to produce DODs, and participants overall were more likely to produce DODs with verbs that were featured in DOD (as opposed to PD) sentences in the bias phase. Participants were also more likely to produce DODs with DOD-biased verbs. Despite the numerically larger likelihood of structure repetition after matching as opposed to mis-matching sentences, neither prime bias match (β= 0.22, $\chi 2$ (1) = 2.21, p = .12) nor its interaction with prime structure (β= -0.22, $\chi 2$ (1) = 0.29, p = .59) reached significance, showing that DOD production was not significantly influenced by whether participants heard the structures in a matching or mis-matching sentence in the test-phase. None of the other effects reached significance either (all p values > 0.59). Supplementary analyses with baseline DOD production included yielded the same pattern of results, so are not included here.

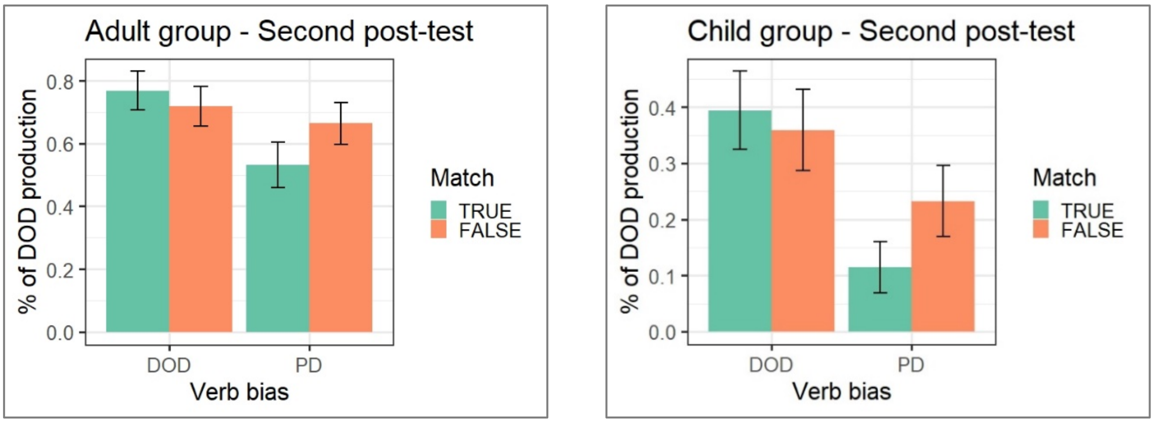

Figure 3.6. – Proportion of DOD production in the second post-test in both age groups. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

3.2.3.3.2. Adult group

The full model including the adult dataset in the second post-test phase included prime-structure, verb-bias match and target verb bias as fixed effects and by-subject random intercepts. The pattern of results was similar to that found in the previous analyses. The main effects of prime structure (β= 0.34, $\chi 2$ (1) = 7.08, p < .008) and target verb bias (β= 3.19, $\chi 2$ (1) = 13.51, p < .0003) were significant, showing that participants were significantly more likely to produce a DOD structure if they heard the target verb in a DOD structure in the bias-phase and if the target verb was DOD-biased. Neither the prime bias match (β= 1.5, $\chi2$ (1) = 1.74, p =.19) nor its interaction with prime structure (β= -0.14, $\chi2$ (1) = 0.09, p = .76) reached significance in this analysis either.

3.2.3.3.3. Child group

The full model on the child dataset in the second post-test phase included prime-structure, verb-bias match and target verb bias as fixed effects and by-subject random intercepts. The pattern of results did not differ in this age group either. The main effects of prime structure (β= 0.42, $\chi2$ (1) = 6.39, $p < .02$) and target verb bias (β= 3.29, $\chi2$ (1) = 14.89, $p < .0002$) were significant, showing that participants were significantly more likely to produce a DOD structure if they heard the target verb in a DOD structure in the test-phase and if the target verb was DOD-biased. Neither the prime bias match (β= 0.28, $\chi2$ (1) = 0.74, p =.39) nor its interaction with prime structure (β= -0.27, $\chi2$ (1) = 0.21, p = .65) reached significance in this analysis either.

3.2.3.3.4. Summary of the results of the second post-test

While both age groups were more likely to re-use the dative structure they heard the verbs with in the test-phase (while verb-bias was taken into account) this effect was not modulated by how surprising the dative structure in the bias phase was. This study therefore provided no evidence for verb-based error-based learning effects.

3.2.3.4. Summary of the results

Both the child and the adult group showed the predicted pattern in response to our key manipulation of interest: participants produced more DODs in DOD- and PDs in the PD bias group compared to their pre-test performance. This result however was only significant in the full (child and adult) model when including participants who did not show ceiling performance in the pre-test phase, was not significant in the adult-only model and was only significant in the child-only model when syntactic knowledge and age were taken into account. The results did not show the expected patterns in the other phases of the study. In the bias phase, only adult participants showed an immediate prime surprisal effect (when their baseline DOD production was taken into account) and while in the second post-test all participants were more likely to repeat the dative structure that verbs appeared with earlier, the likelihood of structure repetition did not depend on whether the structure was surprising or predictable. In other words, neither the child nor the adult group showed verb-dependent error-based learning effects.

Table 3.3.

Appearance of expected response patterns per study phase and age group.

	Age group	
	Adults	**Children**
Phase 2 - **Prime surprisal (priming)**	✓ (✓) Baseline DOD production included, priming n.s.	✗ (✓)
Phase 3 - **Abstract learning**	✓ n.s.	✓ No ceiling in pre-test, age and TROG included

Phase 4 - **Verb-based learning (structure repetition)**	✗(✓)	✗(✓)

In Phase 2 the table shows whether participants demonstrated immediate prime surprisal and structural priming effects, the latter in brackets. At Phase 3 the table shows whether participants showed more learning for abstract structures after surprising as opposed to predictable sentences. At Phase 4 the table shows whether verb-based learning rates were higher in surprising sentences. In brackets we can see whether participants were likely to use the dative structure in Phase 4 that specific verbs appeared with in the bias phase.

3.2.4. Discussion

The goal of our study was to evaluate the central prediction of error-based theories of language acquisition, that surprising linguistic input leads to higher rates of learning than predictable input. To achieve this, we embedded a prime surprisal study (typically used to compare immediate priming effects after surprising versus predictable sentences) in a four-stage intervention study to assess both the short and longer-term effects of predictability. In the first, pre-test phase of the study, we assessed participants' baseline rates of dative production. The main goal of the second, bias-phase was to bias the participants towards one of the dative structures. Here we split participants into a DOD- and a PD-bias group. We always presented one of the structures (e.g. DODs in the DOD-bias group) in surprising sentences by pairing them with verbs that more frequently appear with an alternative structure, while the other structure (PDs in the DOD-bias group) always followed verbs that they frequently appear with, making these structures predictable. We also aimed to replicate immediate prime surprisal effects (e.g. Peter et al., 2015) in this phase. The goal of the crucial post-test phase was to compare participants' rates of dative production depending on which bias-group they were in. We expected that participants' pre- to post-test production would shift towards the dative structure they were exposed to in surprising (as opposed to predictable) sentences in the previous phase. The fourth phase, second post-test, targeted potential verb-specific error-based learning effects. We assessed whether participants reused the dative structure that specific verbs appeared with in the test-phase and whether structure repetition was more likely if the structure appeared in surprising as opposed to predictable sentences.

First, we discuss the results of the post-test phase, where we measured the impact of our key manipulation of interest. As predicted by error-based learning theories, both age groups were more likely to produce DODs in the DOD- as opposed to the PD-bias group. Furthermore, both adult's (average 4.25% shift) and children's (average 6% shift) pre- to post-test production shifted towards the dative structure they were exposed to in surprising sentences in the previous phase. This effect, however, did not reach significance in our registered analyses. This may be due to the differences in the datasets we simulated for our power calculations and the data we collected. First, children's DOD production showed a bigger variance in both the pre-test and post-test phases compared to the simulated data. Second, even though we expected no ceiling performance (allowing for either a positive or negative post-test shift for each participant) 5 children and 18 adults produced only DODs and 9 children and 2 adults produced only PDs in the pre-test phase. As participants with pre-test ceiling performance can only shift towards one of the structures in the post-test phase, the post-test production of these participants was not informative about the impact of this manipulation. Thus, we carried out an additional analysis excluding participants with ceiling pre-test performance. Despite the lower statistical power (due to the decreased number of trials), this analysis demonstrated significant learning effects in the group as a whole, and in the child group when age and syntactic knowledge were taken into account. To our knowledge, this is the first study that has found such learning effects, providing tentative experimental evidence for a central claim of error-based learning theories: that surprising input leads to more learning compared to predictable input. These results are in line with previous studies demonstrating that children's production frequencies can be shifted towards a less frequent structure by exposure in the bias phase (e.g. Branigan & Messenger, 2016; Kidd, 2012), and contributes to this existing literature by specifically showing that these differences are related to input predictability.

These results also speak to another important claim of the dual-path model: that learning effects decrease as the learner accumulates more linguistic knowledge. Our study addressed this question in two ways. First, we compared learning effects in the child and adult group. While children showed a numerically larger pre- to post-test shift than adults, the interaction of age group and bias-dependent post-test performance did not reach significance. The reason for this may be that our study was underpowered for this contrast since we did not model it in our power calculations. Second, we assessed the contribution of age and syntactic knowledge (measured by the TROG test) to the learning effects in the child group. While TROG score had no significant effect, we found that younger children were more sensitive to the intervention. The fact that we found an effect of age is in line with the Dual-path model, which predicts that syntactic representations become stronger as the learner accumulates more knowledge, and these representations are then less sensitive to modification. The fact that we did not find an effect of TROG score is not in line with the predictions of the Dual Path model, though it is possible that the TROG test was not sensitive enough regarding the aspects of syntactic knowledge that influence error-based learning in this task. Either way, it is important to note that the above effects did not reach significance in the pre-registered analyses, so they need to be replicated in follow-up studies that are adequately powered for these comparisons.

Second, we discuss the results of the bias-phase. Numerically, we replicated the prime surprisal effects found in previous studies, but only in the adult group. As predicted, the priming effects in the adult group were larger after surprising as opposed to predictable primes in our study, though neither the priming nor the prime surprisal effect reached significance in the pre-registered analyses. However, while we found a significant priming effect in the child group, there was no sign of prime surprisal. In fact, children were slightly (but not significantly) less likely to repeat the prime structure when it appeared in an expected as opposed to unexpected sentence structure.

In order to discuss possible explanations of these effects, we first need to consider the differences between our bias phase and other studies targeting immediate prime surprisal effects. Since the main goal of this phase was to induce surprisal-based learning effects, we divided participants into a DOD-bias and a PD-bias group. In the DOD-bias group participants always heard DOD sentences in a surprising sentence context and PDs in a predictable one, while participants in the PD-bias group were subjected to the opposite pattern. This created a partial between-subjects design where participants were either subjected to only DOD-biased verbs (paired with either DOD (predictable) or PD (surprising) structures) or only PD-biased verbs (paired with either PD (predictable) or DOD (surprising) structures), see Table 3.2.). In previous prime surprisal studies all participants heard all four types of sentences. While this design was necessary to contrast learning rates for predictable versus surprising sentences, it may have interfered with the immediate prime surprisal effects for a number of reasons.

First, as each participant was exposed to only two out of four conditions, the differing baseline DOD-production rates may have had a larger influence on the immediate prime surprisal effects than in a fully within-participant design. Second, in other prime surprisal studies participants heard all four sentence types in a randomized order, and therefore had roughly equal cumulative experience with sentences that could bias them towards the DOD- or the PD-structure throughout the experiment. This equal cumulative experience would be unlikely to interfere with the immediate prime surprisal effects. As the main goal of our bias phase was to bias participants towards one of the structures, they only heard two out of four conditions thus their cumulative experience throughout the test-phase was already skewed towards either the DOD or the PD structure. It is then possible that these biased cumulative effects have interfered with the immediate prime surprisal effects. Third, to assess verb-based, as well as abstract learning effects, we consistently paired each verb with one of the dative structures for each participant (e.g. *give* only appeared with DODs for participant 1 in the test-phase). This meant that since each participant was exposed to each verb-structure pairing four times, these combinations became more and more predictable throughout the test-phase and may not have been surprising enough to induce immediate prime surprisal effects at the later stages of this phase.

To assess whether the above differences interfered with the prime surprisal effects, we carried out three exploratory analyses on the test-phase production of both age groups. In the first analysis we included participants' baseline DOD-production (as measured in the baseline-phase of our study), in the second analysis we assessed whether participants were overall more likely to produce DODs in the bias phase if they were assigned to the DOD- as opposed to the PD-bias group and in the third analysis we targeted immediate prime surprisal effects by only including the first two occurrences of each verb-structure pairing. Adults showed significant prime surprisal effects when their baseline DOD-production rate was taken into account, but there was no significant prime surprisal effect in the child group. Furthermore, the exploratory analyses did not provide evidence for the overall DOD- or PD-biased bias-phase biasing the immediate prime surprisal results, nor did they suggest that the increasing predictability of verb-structure pairings inhibited prime surprisal effects. Thus, none of the exploratory analyses could explain the lack of immediate prime surprisal effects in the child group.

The last phase of the study, the second post-test, targeted verb-dependent error-based learning effects. In this phase we expected participants to be more likely to use the same dative structure that specific verbs appeared with in the bias phase. We also expected that the likelihood of structure repetition would be higher if the structure was unexpected in the bias phase. While we found that participants in both age groups were significantly more likely to reuse the structures the verb appeared with previously, this effect was not modulated by how surprising the structure was. This study therefore does not provide evidence for verb-dependent error-based learning effects.

However, we are reluctant to draw strong conclusions based on the absence of these effects in the current work. This phase of the study was exploratory and provided a less strong test of learning than the main test of abstract learning. The partially between-participant design led to both uneven target verb-bias rates and uneven baseline DOD-rates in the different conditions that may have masked the learning effects. Furthermore, both participants' abstract learning effects and their previous dative production with the same verbs may have interfered with the results in this phase.

Overall, we have replicated prime surprisal effects in the adult (but not in the child) group and, crucially, shown that sentences that are more surprising led to more learning than predictable sentences. We have also shown that these learning effects are stronger in the younger children. These results are mostly in line with the predictions of error-based learning theories. However, as our key comparison of interest has only yielded significant results in exploratory analyses, the replication of this effect using more accurate power calculations is crucial.

In addition, some of our results are inconsistent with the predictions of error-based learning theories. For example, error-based learning theories would predict verb-based effects in the second post-test. We did not find such effects though they could have been masked or compromised by multiple factors, which means we cannot draw strong conclusion based on their absence. The most challenging result for an error-based leaning account, though, is the presence of longer-term learning effects without immediate prime surprisal in the child group. The Dual-path model suggests that immediate priming effects are the product of the same learning mechanism that lead to long-lasting changes in syntactic knowledge. It would thus predict similar effects of immediate prime surprisal and learning. The disconnect between immediate priming effects and long-term learning in the child group raises questions about whether these effects are always induced by the same mechanism. However, once again, as the main goal of our study was not to assess the relationship between immediate prime surprisal and long-term learning, the results of this comparison must be interpreted with caution.

3.2.4.1. Limitations.

The study had a number of limitations. First, while the different stages of the study addressed different aspects of language change, the main target of our study was abstract syntactic learning based on surprising versus predictable input. In order to ensure that the potential effect of our main manipulation was due to the different levels of predictability of the sentences in the test-phase, we had to make compromises when designing the other phases targeting immediate prime surprisal and verb-based learning. These compromises may have interfered with some of our results. We have discussed these modifications and their potential consequences in the previous sections.

Second, since, to our knowledge, this is the first study directly targeting learning effects based on surprising versus predictable input, we had to base our power calculations on studies targeting related (but not identical) effects. This made it especially challenging to realistically model our outcomes. While the effect sizes in the child group were close to our original estimates, our power calculations did not manage to take into account some important factors, such as larger than expected variance in DOD-production and interference from at ceiling participants. As a result, although our results overall support the main hypothesis tested, it is crucial to replicate these results based on a power calculation taking the above factors into account.

The final limitation of our study lies in the nature of our method, the prime surprisal paradigm. While it can directly address the potential changes in language production depending on the predictability of the input, it only takes measurements at the beginning and the end of the learning event (when the predictable versus surprising input is heard and when the subsequent production occurs). It does not provide information about the progression of the learning events. It is therefore crucial to pair this method with on-line measures such as EEG or eye-tracking to gain more information about the details of how these learning effects unfold over time.

Overall this study did not obtain clear-cut results for many of the questions of potential interest. This is largely due to the exploratory nature of the current project: we both addressed a question which has not yet been directly targeted in the literature and used a novel experimental design. As such, planning for all contingencies before data collection was especially challenging. Furthermore, the decision to prioritize exploring the main abstract learning effect in the experimental design led to less conclusive secondary results (concerning immediate prime surprisal and verb-based learning). However, it did mean that we could ensure that the differences in our main comparison assessing abstract learning rates are due to different levels of predictability. This project has also shown that the current paradigm is effective in targeting error-based learning in children, and can now be adapted to target a range of different effects, offering the opportunity to expand our knowledge of error-based learning theories of language acquisition.

3.2.5. Summary and conclusions

Our study embedded the prime surprisal paradigm in a four-stage intervention study to address a strong, but as yet not directly tested, claim of error-based learning studies that surprising input leads to more learning than predictable input. Our results are broadly consistent with this prediction showing that both adults' and children's dative production shifted towards the structure they were biased towards in the previous phase. This effect was stronger in younger children. However, while the above findings are in support of error-based learning theories, some of our other results are not predicted by them. First, while we found immediate prime surprisal effects in the adult group, they were not present in the child group. As the Dual-path model proposes that immediate priming effects and long-term learning are the product of the same learning mechanism, the existence of longer-term effects in the absence of immediate prime surprisal in the child group is especially challenging for this account. Furthermore, as opposed to the predictions of error-based learning theories, this study did not detect verb-dependent error-based learning effects in either age group. Due to the project's overall exploratory nature and the study's strong focus on abstract learning the additional effects and the overall pattern of results need to be interpreted with caution and require further investigation. To our knowledge this was the first study providing direct evidence for faster learning rates after surprising versus predictable input. As well as providing crucial information about error-based learning theories the present work also contributed to establishing an experimental paradigm that can be used to target further aspects of error-based learning theories of language acquisition in the future.

CHAPTER FOUR: PROCESSING SURPRISING AND PREDICTABLE SENTENCES – EVIDENCE FROM ERPS

4.1. Introduction

Error-based language acquisition theories describe an interactive learning mechanism that works via linguistic prediction (Chang, Dell & Bock, 2006; Ramscar, Dye, & McCauley, 2013). These models suggest that listeners are constantly anticipating the next word in the speech stream, then comparing this prediction to the actual input. If the prediction and the input do not match, an error signal is created. This error-signal is then used to modify linguistic knowledge accordingly. One of these theories, the Dual Path model (Chang et al., 2006), simulates both syntactic acquisition and lifelong adaptation of syntactic knowledge. It proposes that the learning mechanism that leads to language acquisition in childhood stays active in adulthood, and aids adaptation to different speakers or linguistic environments. According to this model, processing expected and surprising input influences adults and children similarly, although the effects are weaker in adulthood as more established knowledge is less sensitive to unexpected input.

The Dual Path model is a connectionist model, where frequency-based syntactic knowledge is represented in connections associated with different aspects of knowledge, such as the overall likelihood of the different dative structures (e.g. DOD, double object dative versus PD, prepositional dative) or how often these structures occur with certain verbs (e.g. the co-occurrence between the verb *sell* and the DOD or the PD structure). As the DOD structure is overall more frequent in English, the connections between the nodes that store representations describing transfer events with this structure are stronger than those using the PD structure.

However, as the verb *sell* appears with the PD structure more frequently than with DODs, the connections that represent the link between *sell* and the PD structure are stronger than those that represent the link between *sell* and the DOD. Like other error-based learning theories, the Dual Path model also operates using an error-based learning mechanism in which the processing of unpredicted structures results in the production of an error-signal. This error signal then leads to an increase in the weights of the connections that represent the knowledge of the unpredicted structure. For instance, when the listener hears *The boy sold ...*, they are most likely to predict a PD continuation (*... the ball to the girl*), as PDs follows the verb *sell* more often than DODs. Thus, if they heard a DOD continuation (*...the girl a ball*) instead, an error signal would result from the discrepancy between the expected PD and the DOD that was actually heard. This error signal would then be used to increase the weights of the connections that represent the link between the message and the unexpected DOD structure. This error-based mechanism has various consequences. Crucially for the experimental assessment of error-based learning theories, these adjustments are assumed to lead to the immediately increased likelihood of the speaker repeating the unexpected DOD structure – otherwise known as a priming effect. However, the most important role of these adjustments is that, over time, they accumulate and drive language acquisition by leading to long-lasting changes in syntactic knowledge.

In line with the above prediction, multiple priming studies have found a larger immediate syntactic priming effect after surprising (e.g. DOD structure appearing after a PD biased verb) as opposed to predictable (e.g. DOD structure appearing after a DOD biased verb) prime sentences. In other words, these studies demonstrated a prime surprisal effect. Prime surprisal appears under various conditions: for instance, in comprehension (Fine & Jaeger, 2013), in production (Jaeger & Snider, 2013) and in a language other than English (Dutch; Bernolet & Hartsuiker, 2010). Crucially with regards to demonstrating prediction's role in language acquisition, these effects have also been demonstrated in a dative priming study involving English-speaking child participants. Peter et al.'s (2015) study takes advantage of different dative verb biases: some verbs appear more often in the DOD structure (such as the verbs *give* or *offer*) and some verbs prefer the PD structure (such as the verbs *sell* or *slide*). Using these biases, Peter and colleagues' study contrasted priming effects after predictable sentences where the verb-bias and the sentence structure matched (e.g. DOD structure – DOD biased verb, such as *Boots gave Dora a puppy*.) and surprising sentences where there was a mis-match between the sentence structure and verb bias (e.g. DOD structure – PD-biased verb, such as *Boots gave a puppy to Dora*.). They found that a group of 3-4 and 5-6-year-old children, as well as a group of adults, were more likely to repeat the prime structure if the sentence was surprising, showing that prime surprisal effects can be detected from an early age.

Up until recently, prime surprisal studies concentrated on immediate effects and did not directly assess the central claim of error-based learning theories: that surprising input leads to more learning than predictable input. However, a new investigation has targeted longer-term learning effects as well: Fazekas, Jessop, Pine and Rowland (in principle acceptance; reported in Chapter 3 of the present work) assessed whether 5-6-year-old children and adults learn more about dative structures when they are surprising (as opposed to predictable) in the given sentence context. They designed a four-phase study combining the prime surprisal method with a paradigm assessing whether the original priming effects are cumulative and persistent (e.g. Kaschak, Kutta & Jones, 2011). The first, pre-test, phase assessed participants' baseline dative production as they described videos depicting transfer actions carried out by familiar characters (such as Dora handing a ball to Boots., Rowland et al., 2012; Peter et al., 2015). The second, bias, phase was designed to induce immediate prime surprisal effects as well as induce a longer-term bias in participants' dative production towards either DOD or PD structures. While participants in both groups heard an equal number of DODs and PDs overall, half of the participants (the DOD-bias group) always heard DODs in a surprising sentence context (with mis-matching verbs) while PDs were featured in predictable sentences (with matching verbs). The other half (the PD-bias group) heard surprising PDs and predictable DODs. The final, post-test, phases of the study assessed whether participants' pre- to post-test performance was more likely to shift towards the dative structure they heard in surprising sentences. According to the Dual Path model, participants in the DOD-bias group would only produce error signals after processing the unexpected DOD structures while participants in the PD-bias group would produce error signals after PDs. These error signals would in turn strengthen the connections between the DOD structure and messages containing transfer actions in the DOD-bias group, and connections with the PD structure in the PD-bias group. These stronger weights would then lead to an increased likelihood of the participants producing the structure they heard in unexpected sentences in the final, post-test phase. In line with the above prediction, Fazekas and colleagues found that participants in the DOD-bias group produced more DODs compared to their pre-test performance, whereas participants in the PD-bias group shifted towards higher PD production. This outcome is in line with the central prediction of error-based language acquisition theories, suggesting that more surprising input leads to more changes in the weights of the connections supporting abstract structures and these changes not only lead to change in short term production (prime surprisal), but also in the longer term (learning).

The prime surprisal paradigm is an excellent method for targeting error-based learning as it assesses a learning event from its starting point (appearance of surprising or predictable input) to its outcome (potential repetition of the input). However, as this method does not assess the steps between the start and endpoint, we cannot determine how processing more surprising sentences leads to an increased likelihood of structure repetition and whether the potential processing differences between predictable and surprising sentences are compatible with error-based learning accounts. For instance, we do not know whether the processing differences appear at the structure decision point as we would expect in case of word-by-word processing suggested by the Dual Path model or later on in the sentence when participants gain more information on how the structure unfolds. We also do not know whether the processing differences are related to active predictions or they stem from an alternative mechanism. To address this question, we took ERP (event-related potential) measurements during a three-phase prime surprisal study (based on Fazekas et al.) to gain on-line information about the processing of predictable and surprising sentences.

ERPs have been widely used to target linguistic predictions. In these studies, participants are typically exposed to linguistic stimuli, and ERP measurements on more or less predictable words are compared (see for instance the sentence-final words of 1.a. and 1.b. from Kutas & Hillyard, 1984). These studies identified different ERPs, such as the N400 or the P600 that are modulated by stimulus predictability.

1.a. Don't touch the wet **pain**t!

1.b. Don't touch the wet **dog**!

The N400 is a negative-going ERP component that peaks at 300 to 500 ms after word onset, and that has a centro-parietal distribution (Kutas & Hillyard, 1983). It is sensitive to ease of processing, and is modulated by various aspects of language, such as word frequency, neighbourhood size and contextual predictability (e.g. Kutas & Hillyard, 1980, Holcomb, Grainger, & O'Rourke, 2002; Barber, Vergara, & Carreiras, 2004). Contextual predictability is typically determined by cloze probability, which is defined by the percentage of participants who chose a given continuation for a sentence stem in an offline task. For instance, if 40 out of 50 people chose *paint* as the best continuation of the sentence 1.a., the cloze value of *paint* would be 80%. Kutas and Hillyard (1984) observed a roughly linear reduction in N400 amplitude as predictability (measured in cloze values) increased and suggested that the N400 might be a marker for linguistic prediction.

Another potentially prediction-related component is the P600, which is a later, positive-going component, peaking 500-700 ms after stimulus onset, and is maximal over posterior electrode sites. The amplitude of the P600 is greater when a stimulus requires additional syntactic processing, such as in garden path sentences, when participants are expecting a certain sentence continuation, but the incoming stimulus disconfirms this and a new interpretation needs to be constructed. For instance, in sentence 2.b. participants tend to first integrate *the house* and *the garage* into one conjoined noun phrase, but when they hear the word *is*, they need to reassess their previous interpretation and construct a new interpretation instead, where *the garage* is the start of a new clause (Kaan & Swaab, 2003). The reassessment in sentences similar to 2.b. typically results in an increased P600 compared to unambiguous sentences, such as 2.a.

2.a. The man is painting the **house but the garage is** already finished.

2.b. The man is painting the **house and the garage is** already finished.

P600 effects are not exclusively elicited by differences in syntactic processing. Kuperberg and colleagues (Kuperberg, Sitnikova, Caplan & Holcomb, 2003) found larger P600 effects when the upcoming verb was likely to lead to thematic role re-assignment, as in sentence 3.a. versus 3.b. In sentence 3.a. participants are most likely to assign the theme role to the inanimate *eggs* up until the appearance of the verb, which then suggest that *eggs* is in fact the agent of the sentence. Sentence 3.b. is less likely to elicit such thematic role re-assignment as participants are likely to assign the agent role to the animate *boys* early on and this role assignment is not challenged by the verb.

3.a. For breakfast the **eggs** would only **eat** toast and jam.

3.b. For breakfast the **boys** would only **eat** toast and jam.

The goal of our study was to use EEG to evaluate whether the processing differences between predictable and surprising sentences are compatible with those suggested by error-based learning theories. To assess this, we took ERP measurements on dative sentences that were similar to those that led to prime surprisal and learning effects in previous studies (e.g. Peter et al., 2015, Fazekas et. al, in principle acceptance, Chapter 3), see sentences 4.a. to 4.d. We also recorded participants' sentence production to assess whether the aforementioned immediate prime surprisal and surprisal-based learning effects could be replicated in the present ERP study.

4.a. Lisa gave the squirrel a hedgehog.

4.b. Lisa gave the hedgehog to the squirrel.

4.c. Lisa sold the squirrel a hedgehog.

4.d. Lisa sold the hedgehog to the squirrel.

Our study targeted the sentence region where error-based learning accounts suggest processing differences due to predictions being confirmed or disconfirmed. This region is at the structure decision point, where it becomes clear to the participants whether they are processing a DOD or a PD structure and whether this structure matches their expectations or not. Without supporting context, the first difference between PD and DOD sentences is at the words *to* or *a* (see orange words in sentence 4.a. to 4.d.). However, if the listener knows the identity of the theme and recipient of these sentences, the structure decision point can move ahead to the second noun of the sentence (see green words in sentence 4.a. to 4.d.). For instance, in sentence 4.a. when the listener hears *Lisa gave the squirrel* ..., knowing that the squirrel is the recipient of the sentence, they can identify the structure as a DOD, as in a PD sentence they would have heard the theme (*hedgehog*) instead of the recipient (*squirrel*). Due to the difficulties associated with measuring ERPs on short words (such as *to* or *a*), we included a disambiguating picture before the sentence stimuli began to play (See Figure 4.1. in Section 4.2.5), to make participants aware of the content of these sentences. Thus, in our paradigm, participants always saw a picture first that identifies the agent and the patient of the action as well as the target verb. Then they heard a sentence, but the sentence presentation only started once the picture disappeared from the screen. As by the sentence start participants are aware of the identity of the theme and the recipient, the structure decision point in our sentences always fell on the second noun.

According to error-based learning theories, as sentences are processed word by word, after each processed word a prediction is made for the upcoming word. For instance, in sentences 3.a. and 3.b. above, as listeners process the verb *give*, they start predicting the next word of a dative structure. As *give* is a DOD-biased verb, and participants are aware of the content of the sentence, they would be most likely to anticipate a DOD structure's next words: *the squirrel*, whereas if they were listening to sentences 3.c. and 3.d. which use the PD-biased *sell*, they would start anticipating the next words of a PD structure: *the hedgehog*. Once the actual input appears, participants assess whether their prediction matches the input and if there is a discrepancy, an error signal is created. For instance, when participants are listening to sentence 3.a. or 3.b. and expect to hear *the squirrel*, after the DOD-biased *give*, their prediction is confirmed in 3.a. but disconfirmed in 3.b. This disconfirmed prediction could lead to immediate processing differences that are detectable with ERPs. Moreover, according to error-based learning theories they also result in the creation of an error signal that leads to both an increased likelihood of immediate structure repetition as well as long-term learning. Based on this account, we would expect the most crucial processing differences in these sentences to occur during the processing of the second noun. Therefore, this location was the target of our examination.

We analysed two time-windows in this location, one from 300-500 ms and one from 500-700 ms, to target potential differences in the N400 and P600 components, which have been previously shown to be sensitive to predictability. As the Dual Path model suggests that participants had already predicted the second noun by the time they heard it in the predictable sentences (4.a. and 4.d.) but not in the surprising ones (4.b. and 4.c.), we expect a larger N400 in the surprising sentences. Furthermore, a mistaken prediction would mean that participants would have to construct a new structure instead of the one they have been predicting up until that point so we would also expect larger P600 effects indicating increased processing costs during the surprising sentences. For instance, in 4.b. participants are most likely to start constructing a DOD structure after hearing the DOD-biased verb, but when they hear *hedgehog* indicating a PD structure they need to abandon the DOD structure and start constructing a PD instead. This would lead to a larger P600 effect indicating increased syntactic processing.

Overall, we took three behavioural and ERP measures in this study. We compared both the immediate and the longer-term effect of processing surprising and predictable sentences and also took ERP-measures at the structure decision points of these sentences.

4.2. Methods

4.2.1. Study summary

The aim of this study was to learn more about the role of prediction in syntactic processing and production in both the short and longer term. The first goal of the study was to assess the potential differences in the neural correlates associated with processing expected and surprising continuations of sentences as a function of verb bias. Our second goal was to replicate earlier results assessing the main prediction of error-based learning theories that less predictable (more surprising) linguistic input leads to both more immediate and more lasting language change than more predictable input. To achieve the above goals, we conducted a three-phase experiment designed to induce error-based learning via prime surprisal while recording EEG during sentence processing.

4.2.2. Study design

The prime surprisal paradigm used in this study capitalises on the fact that some verbs are substantially more likely to appear in one dative sentence structure than another in English, and are thus surprising, despite being grammatical, in the alternative structure. Error-based learning predicts that there will be a bigger change in syntactic representations (i.e. learning) after surprising (e.g. DOD structure occurring after a PD biased verb) than unsurprising (e.g. DOD structure occurring after a DOD biased verb) primes. Learning is defined as a change in the underlying syntactic representations and is operationalised as a performance difference in the post-test phase of the study depending on which bias group the participants were assigned to. More specifically, learning will be deemed to have occurred if the participants are significantly more likely to produce DODs in the post-test phase of the DOD-bias rather than the PD-bias group when their baseline DOD-production rate (measured in the pre-test) is accounted for.

In the first, baseline phase of the study, we assessed participants' baseline rates of dative production (i.e. how many DODs and PDs they produced). Participants described target pictures depicting transfer actions involving three participants (X was transferred from Y to Z) and were free to choose either PD or DOD structures.

The second, priming (or bias) phase was designed to bias the participants towards one of the dative structures while eliciting immediate prime surprisal effects. Here, participants described pictures depicting transfer actions in a similar way to the baseline phase but they also heard prime sentences containing either a DOD or a PD structure. Both structures were consistently paired with either matching or mismatching verbs in the prime sentences for each participant (e.g. DODs only appeared with mis-matching verbs, while PDs only appeared with matching verbs for group A and vice versa for group B). This way, participants in group A always heard PDs in predictable sentences and DODs in surprising sentences while participants in group B heard DODs in predictable and PDs in surprising sentences. The third, post-test phase was similar to the baseline phase, but the goal was to reassess participants' rates of dative production. If less predictable input leads to more lasting language change than more predictable input (as suggested by error-based learning theories), participants in this phase should be more likely to produce the dative structure they were exposed to with a mismatching verb in the bias phase (i.e. DODs for participants in group A) than the structure they heard with a matching verb (PDs in group B). In order to eliminate the influence of lexically-based long-term priming effects, we used different verbs in all phases of the study.

We recorded EEG to assess the potential ERP differences during the auditory processing of predictable and unpredictable prime sentences during the bias phase.

Table 4.1.
General study design showing different trials and verb biases in each phase.

	Group A		Group B	
	Structure	**Verb bias**	**Structure**	**Verb bias**
		Baseline phase		
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
		Test (bias) phase		
Experimenter	**DOD**	**PD-biased**	**DOD**	**DOD-biased**
Participant	Dative	PD-biased	Dative	PD-biased
Experimenter	**PD**	**PD -biased**	**PD**	**DOD-biased**
Participant	Dative	DOD-biased	Dative	DOD-biased
Experimenter	**DOD**	**PD-biased**	**DOD**	**DOD-biased**
Participant	Dative	PD-biased	Dative	PD-biased
Experimenter	**PD**	**PD-biased**	**PD**	**DOD-biased**
Participant	Dative	DOD-biased	Dative	DOD-biased
		Post-test phase		
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased
Experimenter	Filler	NA	Filler	NA
Participant	Dative	Equi-biased	Dative	Equi-biased

4.2.3. Participants

Thirty-five adult monolingual English-speakers participated in the study. They were recruited from the student participation pool of UC Davis. All participants were right-handed, monolingual native English speakers who had no history of language issues or neurological impairments. All participants signed an IRB UC Davis approved consent form. We excluded eleven participants: three participants did not complete all 6 blocks of the study, six participants had fewer than 30 trials after artefact rejection and we also excluded two participants due to experimenter mistakes. We included data from 24 participants in our final analyses.

4.2.4. Design

Overall structure bias (DOD-bias or PD-bias) was a between-subject variable and prime-type (DOD and PD), verb-bias match (match (predictable) or mismatch (surprising)) and phase (pre-test, bias phase and post-test) were within-subject variables. The dependent variable was the choice of dative structure in the target trials in the behavioural part of the experiment and EEG data time-locked to the second noun in the prime sentences for the neural part of the study.

4.2.5. Visual stimuli

The study contained pictures featuring cartoon figures. The prime and target pictures always included a human character (e.g. the cook or Mary) and two animals (e.g. giraffe, elephant) and depicted transfer actions that can be described using dative sentences. Each participant saw 400 pictures: 180 pictures depicting transfer actions that can be described with prepositional or double object datives for the prime and target sentences and 220 pictures depicting non-causal actions for the filler sentences.

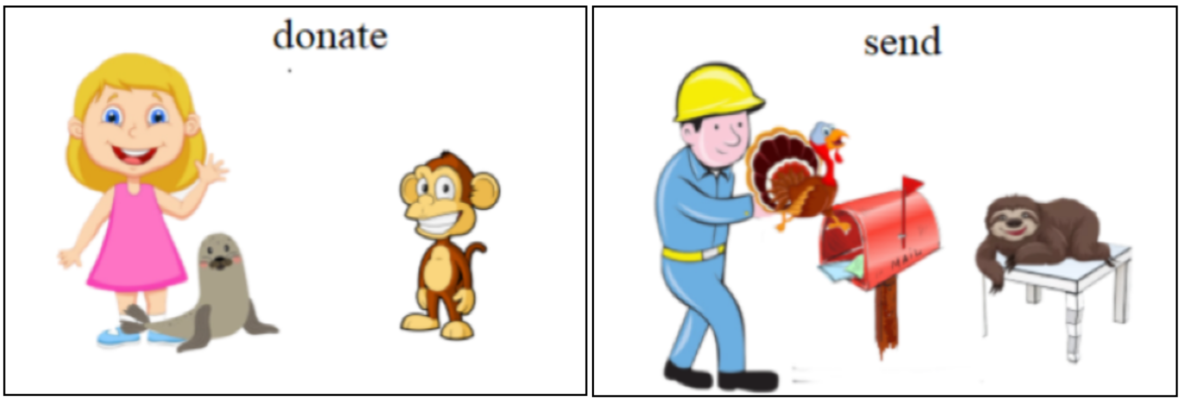

Figure 4.1. Example prime and target picture

The pictures included 80 human characters and 80 animals (please see sentence lists section 6. Appendix for a full list of the characters featured). To avoid effects stemming from experiences with certain character-structure pairings, a different set of characters was used in the prime and target sentences. Furthermore, every animal only appeared once per condition and per role (theme or recipient) for each participant. Human characters only appeared once per condition. In order to control for the possibility that the direction of transfer may influence structure choice, the pictures depicted the direction of motion of transfer actions equally often from right-to-left and from left-to-right.

4.2.6. Sentence stimuli

Each participant heard 400 sentences (including 100 verb stems): 80 prime sentences and 80 target stems plus 80 complete (prime-type) and 80 sentence stem (target-type) fillers in the bias phase and 10 target stems and 20 complete and 10 sentence stem filler sentences in the pre- and post-test phases. Half of the prime sentences appeared as DOD sentences and half as PD sentences. Both structures were consistently paired with either matching or mismatching verbs in the prime sentences for each participant (e.g., PDs only appeared with matching verbs while DODs only appeared with mismatching verbs for participant A and vice versa for participant B). The target sentences were produced by the participant (as either DOD or PD sentences) based on the pictures and sentence stems. For instance, a prime-target trial in the bias phase included a prime sentence such as *The king gave the friendly cat to a horse* (PD) or *The king gave the friendly horse a cat* (DOD) and participants completed a sentence stem such as *Lisa brought…* as a target sentence. All prime sentences included the adjective *friendly* before the critical noun to provide a comparable baseline for the critical ERP measurements.

In order to avoid lexically-based long-term priming effects, we used a different set of verbs in all three phases. We used three sets of verbs that we selected based on a sentence completion and a forced choice norming study completed by UC Davis students (for more details about the norming study, see section 2.2.3.4. in Chapter 2). The verbs are featured here with their DOD frequencies (per all dative completions) in brackets based on our sentence completion norming study. The first two sets of verbs were used in the pre- or post-test phases and the verbs were selected so that that the overall verb-bias of each phase was around 50%. As the goal of these phases was to measure un-biased dative production we wanted to ensure that the target verbs themselves did not create a strong bias and were comparable between the two sets. Each set contained five verbs. Set 1: *assign* (34%), *mail* (40%), *throw* (43%), *loan* (48%) and *hand* (54%); Set 2: *fax* (35%), *lend* (45%), *show* (47%), *bring* (54%) and *award* (55%). The third set of verbs was featured in the bias phase. This set contains eight PD-biased verbs: *recommend* (16%), *donate* (16%), *present* (24%), *sell* (29%), *deliver* (30%), *pass* (32%), *toss* (36%) and *send* (37%) and eight DOD-biased verbs: *issue* (56%), *slide* (57%), *provide* (58%), *give* (61%), *slip* (61%), *serve* (66%), *offer* (70%) and *promise* (78%).

Semi-randomised stimulus lists were created in which the prime and target sentences always followed each other in the bias phase and the same verb or animal did not appear twice in immediate succession. Furthermore, animals in the critical second noun slot in the prime sentences always appeared in one of the two different halves of the lists and at least 10 sentences apart from each other. To control for sentence-specific preferences, we created eight counterbalance groups to ensure that 1. if the DOD structure consistently appeared with matching verbs in one counterbalance group, it appeared with mismatching verbs in the other (and *vice-versa* for the PD structure), 2. if a target sentence appeared in the pre-test in one counterbalance group, it appeared in the post-test in the other and 3. if a target animal appeared in the first half of the list for one counterbalance group it appeared in the other half for the other. In the test- or bias-phase there was always a pair of filler sentences after every target sentence. In the pre- and post-test phases, target-filler pairs and filler-filler pairs alternated with each other.

4.2.7. Procedure

During the study, participants were seated in an electrically-shielded, sound-attenuated booth. Stimuli were presented on an LCD monitor at a distance of 90 cm. The participants were instructed to look at pictures and describe them by completing pre-recoded sentence stems or to judge whether the pre-recorded picture descriptions match the pictures. Stimuli were presented using Presentation software (Neurobehavioral Systems).

Each trial started with the 1000 ms presentation of a white fixation cross (1.5 cm x 1.5 cm) that appeared at the centre of the computer screen against a black background. After the fixation cross disappeared, a picture appeared for 4000 ms for the experimental and 2000 ms for the filler trials. After picture offset, the fixation cross reappeared on the screen. Then, 1000 ms after the onset of the fixation cross, the audio recording of a sentence (in case of prime trials) or a sentence stem (for targets) was presented. The fixation cross remained on the screen throughout the sentence presentation. Once the sentence presentation was complete, either a checkmark or three dots appeared on the screen.

A checkmark (indicating that the participants had no further tasks) appeared on the screen after the prime trials and prime-type filler sentences and it remained visible until the participants indicated (by button press) that they were ready for the next sentence to start. Three dots appeared on the screen after each target sentence stem or target-type filler, indicating that the participant was expected to finish the sentence by talking into a dictaphone that was placed in front of them. The dots remained visible until the participant indicated (by pressing a button) that they had finished talking and were ready for the next trial. To make sure that the participants were attending to the prime sentences, after 30 of the complete filler sentences an additional written sentence appeared on the screen instead of the checkmark (e.g. “The mole and the zebra were hiking.”). Half of these sentences matched the previously heard filler sentence while the other half did not. Participants were asked to judge whether the sentence matched the previously seen picture and give a response via keyboard press: “true” (“z” on the keyboard) or “false” (“m” on the keyboard).

To minimize ocular and movement related artefacts in the EEG signal, participants were asked to keep their eyes fixated on the cross and to refrain from making eye movements or any other posture changes while it was present on the screen. They were also asked to make sure that they take a break if needed and make all necessary posture adjustments while the checkmark or the dots were on the screen and no audio was playing and get ready for the next sentence before they pressed the button indicating that they had completed the trial.

Participants completed a practice session before the beginning of the experiment. The practice session was set up in the same way as the experimental blocks and contained 20 filler-filler pairs. The practice sentences introduced all animals used in the experimental trials. Animals featured in prime trials were always featured with the adjective *friendly* to familiarize participants with these constructions.

The study consisted of six experimental blocks with breaks in between, during which the participant was offered some refreshments, was allowed to make some posture adjustments while remaining seated and could rest their eyes for as long as they wanted.

4.2.8. EEG Recording and Data Processing

EEG was recorded from 29 tin electrodes, embedded in an elastic cap (Electro-Cap International; Eaton, OH). Horizontal and vertical electro-oculograms were also recorded to monitor eye movements and blinks. All electrode impedances were kept below 5 kOhm. The EEG signal was amplified using a Synamps Model 8050 Amplifier (band-pass 0.05-100 Hz) and digitally recorded at a sampling rate of 250 Hz. Initially all channels were referenced to the right mastoid and later re-referenced to the average of the right and left mastoids.

Data processing and analyses were performed offline using EEGLAB toolbox (Delorme & Makeig, 2004) with the ERPLAB plugin (Lopez-Calderon & Luck, 2014). Independent component analysis (ICA) was used to isolate and remove EEG artefacts due to blinks and saccades (Delorme & Makeig, 2004). After ICA was performed and prior to off-line averaging, all single-trial waveforms were screened for amplifier blocking, muscle artefacts, horizontal eye movements, and any remaining blinks over epochs of 1,200 ms, starting 200 ms before the onset of the critical words. Baseline correction was performed using the −200 to 0 ms prestimulus period.

We only included participants in our final analyses who had at least 30 out of 40 artefact-free trials in the critical second noun slot in all conditions. On average, 7% of trials were excluded following artefact rejection (range: 0-22.5%). Overall we included an average of 37 out of 40 artefact-free trials per condition. Average ERPs were computed over all artefact-free trials for each condition and participant. All ERPs were filtered with a Gaussian low-pass filter with a 25 Hz half-amplitude cut-off. Statistical analyses were conducted on the filtered data. We analysed the second noun in each prime sentence.

4.2.9. Coding of the target sentence productions

The target sentences produced by the participants were recorded on an HccToo digital voice recorder and transcribed and coded off-line as a 'DOD structure', a 'PD structure' or an 'other response'. A target response was considered a DOD if it contained two consecutive noun phrases (e.g. *Lisa gave the friendly … hippo a cat.*), and a PD if it contained a noun phrase and a prepositional phrase headed by *to* (e.g. *Lisa gave a friendly … cat to the hippo.*). Responses were coded as 'other response' if the target sentence could not be classified as a DOD or PD response based on the above criteria (e.g. target responses containing a preposition other than *to* or *The king threw the friendly rat at the mole.*) or incomplete datives such as *Lisa sent the friendly cat*.

4.2 Results

The goal of this study was to investigate error-based learning by mapping the neural correlates of processing more or less predictable sentences, as well as assessing how they influence subsequent language production in the short as well as the long term. We will first discuss the behavioural results of the study (i.e., the utterances produced by the participants) which attempted to replicate immediate prime surprisal effects (e.g. Jaeger & Snider, 2007, Peter et al., 2015) as well as contrasting the longer-term effects of processing more or less predictable sentences (Fazekas et al., in principle acceptance, Chapter 3).

4.2.1. Behavioural results

The target sentences from the test and post-test phases were analysed separately by fitting maximal logistic mixed effect models (Baayen, Davidson & Bates, 2008; Jaeger, 2008) using the lme4 package in R (R core team, 2012, version 3.6.1). Model comparisons were carried out to obtain likelihood ratios and p-values on maximal models where the random effect structure was only simplified if the model did not converge (Barr, Levy, Scheepers & Tily, 2013). No main effects were removed due to non-significance. The dependent measure was the production of the different dative structures (DOD=1, PD=0). All factors were effect/sum coded (Wendorf, 2004).

4.3.1.1. Bias phase

We carried out the first set of analyses on the target sentences produced by the participants in the bias-phase of the study. The figure below contains the mean proportion of PD responses after PD and DOD primes involving either matching or mis-matching verbs. This analysis assessed whether structural priming (a larger proportion of DOD responses after DOD rather than PD primes) and immediate prime surprisal effects (larger priming effects if the prime verb's bias did not match the prime structure) are present in this dataset.

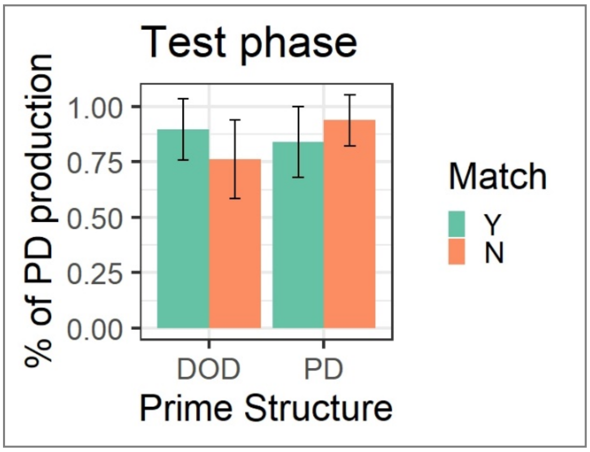

Figure 4.2. Percentage of PD responses after matching and mis-matching DOD and PD primes. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

Our model included (a) prime structure (DOD or PD prime) and (b) prime-bias match (depending on whether the prime structure matched the prime verb's bias) as fixed effects and by-participant and by-item random slopes. The results revealed an overall structural priming effect (participants produced significantly more DOD structures after DOD compared to after PD primes, β= 0.29, $\chi2$ (1) = 7.02, p < .009) although the main effect of match did not reach significance (β= 0.007, $\chi2$ (1) = 0.04, p = .84). Importantly, though, there was a prime surprisal effect manifesting in the significant interaction of prime structure and bias match showing that participants were more likely to repeat the prime structure in the target sentences after mis-matching than after matching sentences (β= 0.66, $\chi2$ (1) = 5.81, p < .02).

4.3.1.2. Post-test phase

The second set of analyses was carried out on the target sentences produced by the participants in the post-test phase. Due to an experimenter mistake, the post-test phase did not contain the appropriate sentence stimuli in the case of two participants.

Therefore this analysis was carried out on data from 22 participants (12 in the DOD-bias and 10 in the PD-bias group). Here we assessed whether participants were more likely to produce more DODs in the DOD-bias than in the PD-bias group, while taking into account their baseline DOD-production, measured in the pre-test phase. Note that in the DOD-bias group participants always heard DODs in surprising and PDs in predictable sentences, while participants in the PD-bias group heard surprising PDs and predictable DODs.

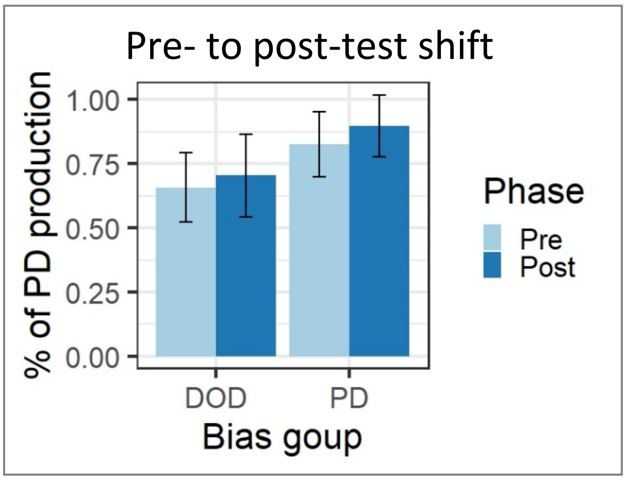

Figure 4.3. Percentage of PD responses in the pre- and post-test phase of the study. (Error bars represent standard error and were generated with remef (Hohenstein & Kliegl, 2013).)

Our model included (a) bias group (DOD or PD-bias group) and (b) baseline DOD production (proportion of the participant's DOD per dative production in the pre-test) as fixed effects and by-participant and by-item random slopes. The results revealed that while participants were 19.26% more likely to produce DODs in the post-test phase of the study if they were assigned to the DOD- rather than the PD-bias group (compared to their baseline DOD performance: 34.3% in the DOD and 17.44% in the PD bias group) the main effect of bias group did not reach significance

(β= 1.16, $\chi2$ (1) = 3.16, p = .075).

4.3.1.3. Summary of the behavioural results

The analyses carried out on the target sentences of the test and post-test phases of the study indicate that participants dative production was more strongly influenced by surprising as opposed to predictable sentences both immediately (prime surprisal) and in the longer term (learning). However the longer-term effects did not reach significance.

4.3.2. ERP results

Our analyses focused on the second noun in the prime sentences, which is the structure decision point in the current study, since the participants also saw pictures of the events (e.g. Emma gave the friendly **horse** a parrot, Emma gave the friendly **parrot** to a horse). At this point, it becomes clear to the listener whether they are hearing a DOD or a PD structure and whether this structure fits their predictions or not. We conducted separate repeated measures ANOVAs using the EZAnova package in R for each condition targeting the two ERP components of interest: the N400 (dependent measure: mean amplitude 300-500ms post word-onset) and the P600 (dependent measure: mean amplitude 500-700ms post word-onset). Significant interactions were followed-up by simple effects comparisons. A Greenhouse-Geisser correction was used for all F tests with more than one degree of freedom in the numerator.

To examine the topographic distribution of the effects, we conducted two analyses: a Midline analysis that included Electrode as a variable (5 levels: Afz, Fz, Cz, Pz, POz), and a Lateral analysis that included the Hemisphere (2 levels: left, right) and Anteriority (2 levels: anterior, posterior) as within-subject variables resulting in four scalp regions, including data from the following electrode groups: left anterior: F7, F3, FC5, FC1; left posterior: CP5, CP1, T5, P3; right anterior: F4, F8, FC2, FC6; right posterior: CP2, CP6, P4, T6. In addition to the aforementioned topographical predictors these analyses included prime structure and verb-bias match as within-participant variables. Details of the analyses can be found below.

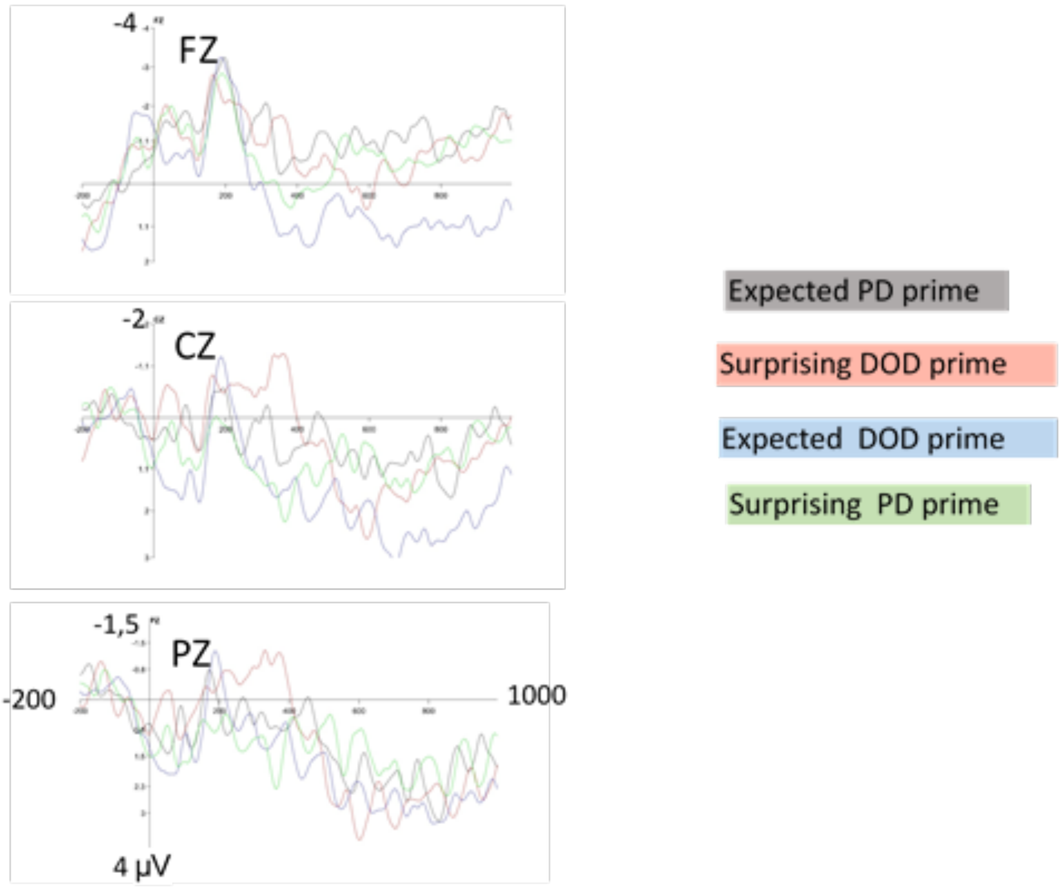

Figure 4.4. ERPs to surprising and predictable sentences in both the DOD and PD prime sentences. Waveforms are plotted for three midline electrodes (Fz, Cz, Pz).

4.3.2.1 N400 window

The repeated measures ANOVAs did not reveal a significant effect of either prime structure ($F(1,23) = 6.04$; $p > 0.9$) or bias-match ($F(1,23) = 2.26$; $p > 0.6$) in the midline analyses. The lateral analyses uncovered a significant three-way interaction between hemisphere, prime structure and bias match ($F(1,23) = 6.84$; $p < 0.02$). However the interaction between prime structure and bias match did not reach significance ($F(1,23) = 3.20$; $p < 0.09$). Simple effect comparisons revealed that, while the interaction of prime and match was significant when only including left-side electrodes in the lateral analyses ($F(1,23) = 5.94$; $p < 0.02$) this effect did not reach significance when only including right-side electrodes ($F(1,23) = 1.108$; $p > 0.2$).

We carried out further simple effect comparisons aimed at the interaction of prime structure and verb-bias match in the left-side electrodes. While the N400 effect was stronger in the DOD than in the PD primes, the main effect of prime bias match did not reach significance in either DOD ($F(1,23) = 4.05$; $p <0.06$) or PD primes ($F(1,23) = 0.03$; $p >0.8$) separately.

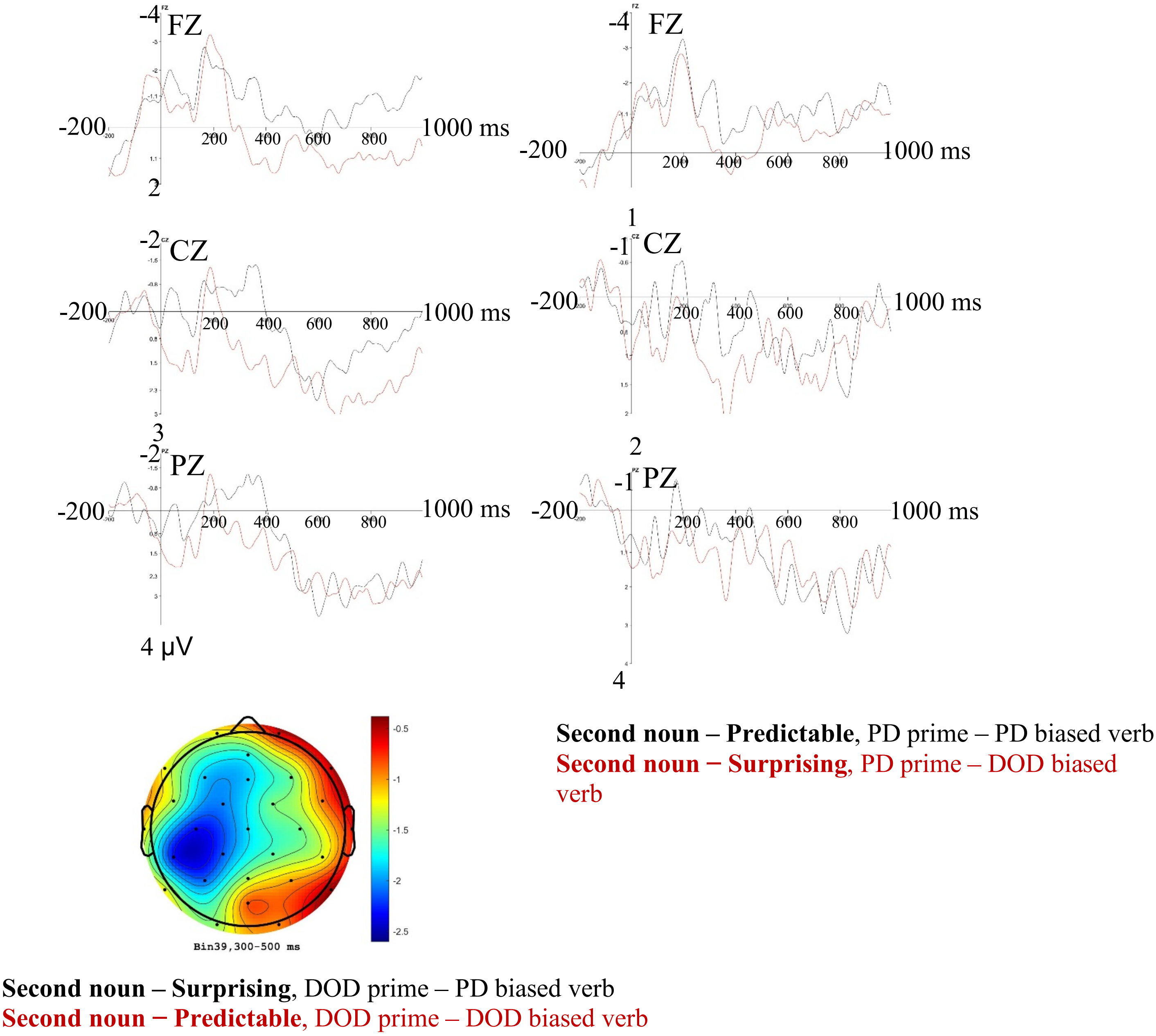

Figure 4.5. ERPs to surprising and predictable DOD (left) and PD (right) prime sentences. Waveforms are plotted for three midline electrodes (Fz, Cz, Pz). The topographic map at the bottom left of the figure shows the difference in mean amplitude between the expected and surprising DOD conditions in the N400 time window for all electrodes.

4.3.2.2. P600 window

The repeated measures ANOVAs did not reveal a significant effect of either prime structure (midline: $F(1,23) = 1.22$; $p > 0.2$, lateral: $F(1,23) = 1.403$; $p > 0.2$) or bias-match ($F(1,23) = 0.07$; $p > 0.7$, lateral: $F(1,23) = 0.003$; $p > 0.5$) in either the midline or the lateral analyses.

4.3.2.3. ERP analyses summary

Neither prime structure nor verb-bias match had a significant main effect in any of the time windows, but the lateral analyses revealed a three-way interaction of hemisphere, prime type and prime bias match in the N400 window. As expected follow-up analyses showed that the prime – bias match interaction was only significant in the left side electrodes. The effect of prime bias match was stronger (though not significant on its own) in the DOD than in the PD primes. In other words, hearing verbs in the mismatching structure elicited an N400 effect in the left hemisphere, especially for PD-biased verbs in DOD structures.

4.4 Discussion

The goal of the present study was to evaluate error-based learning theories by assessing the neural correlates associated with processing surprising and predictable sentences as well as their influence on subsequent syntactic production. In a three-stage prime surprisal-based intervention study, we monitored participants ERPs during sentence processing and also compared both their immediate and delayed sentence production after more or less predictable sentences. We used dative sentences similar to those that led to both immediate and longer-term prime surprisal effects in previous studies. The prime surprisal method builds on dative verb biases showing that all verbs have different dative structure preferences and typically appear more often with either DOD or PD structures. Depending on whether the dative structure following the verb matches the verb's bias (e.g. DOD-biased verb - DOD structure) or not (e.g. PD-biased verb - DOD structure) the sentence ending can be more or less predictable. In previous behavioural studies surprising (mismatching) sentences led to a higher likelihood of structure repetition both immediately (e.g. Peter et al., 2015) and in the longer term (Fazekas et. al, in principle acceptance, reported in Chapter 3). However, as the previous prime surprisal studies did not take online measures during sentence processing, we do not yet know what processing differences lead to this increased likelihood of structure repetition, and whether these differences are compatible with error-based learning theories.

We conducted a three-stage intervention study to address these questions. The goal of the first stage was to measure participants' baseline dative production as they described pictures depicting transfer actions that can be described using datives. In the second bias phase we aimed to bias participants towards either the DOD or PD structure by always presenting one of the structures in a surprising and the other one in a predictable sentence context. We also assessed whether we could replicate the immediate prime surprisal effects found in previous studies. Crucially we took ERP measurements during prime sentence processing in the bias phase to evaluate processing differences between predictable and surprising sentences. The final, post-test phase was similar to the baseline phase but here we assessed whether participants are more likely to produce the structure they heard in an unpredictable (as opposed to predictable) sentence context.

We discuss the behavioural results of the study first, beginning with the bias phase. The goal of this phase was two-fold: first, we aimed to bias participants towards one of the alternative dative structures, second, we aimed to replicate immediate prime surprisal effects found in previous studies (e.g. Peter et al., 2015). In line with the outcome of previous research, we found that participants were more likely to produce DODs after DOD (as opposed to PD) primes, thus demonstrating a priming effect. We also found a prime surprisal effect as priming was larger after surprising as opposed to predictable sentences.

Our second goal was to assess whether we could replicate Fazekas and colleagues' results showing that participants are more likely to produce the structure in the post-test that they were exposed to in surprising sentences in the bias phase, thus demonstrating that predictability not only has immediate, but also longer-term effects on syntactic knowledge. They found that adults were overall 4.25%, and children overall 6%, more likely to produce DODs in the DOD- as opposed to the PD-bias group. We found a similar, but numerically larger difference between our bias groups. Participants in the current study were 12.9% more likely to produce DODs in the DOD- as opposed to the PD-bias group. This result, however, did not reach significance when baseline dative production (measured in the pre-test) was taken into account. This is potentially the result of our study being underpowered for this comparison. The power analyses for Fazekas and colleagues' study using a similar design required 72 participants for both the test- and the post-test comparison to have adequate power. As the main target of the current study was the ERP comparison, we only included 24 participants in the study overall of which only 22 were included in this analysis, potentially not allowing for an adequately powered comparison here.

There is also an important difference between the results of the current study and the previous one. Fazekas and colleagues' study not only showed that participants are overall more likely to produce DODs in the DOD- (as opposed to the

PD-) bias group, but also that participants' pre- to post-test dative-production shifted towards the structure they were biased towards. Our study did not replicate this pattern, as participants in both bias groups produced more PDs in the post-test as opposed to the pre-test. However, this shift was smaller in the DOD-bias group. This difference in the pattern of results may be due to differences in the designs of the two studies. While Fazekas and colleagues used the same set of verbs in the pre- and post-test phase, here we used two different sets of verbs. Thus, in the current study, the potential per-verb individual differences made the study less sensitive to the pre- to post-test shift.

Next we turn to the EEG results. As the goal of this study was to assess error-based learning theories, we concentrated on the sentence region where these mechanisms would predict the appearance of an error signal: the structure decision point. Here it becomes clear to the listener whether they are hearing a PD or DOD structure, and they can also assess whether they have correctly predicted this structure or not. According to error-based learning theories, an error signal will be produced here if there is a discrepancy between the predicted and processed structure. In our study, predictions would be more likely to be correct in matching (e.g. DOD-biased verb followed by DOD structure) as opposed to mismatching (PD-biased verb followed by DOD structure) sentences. Therefore, we expect ERP differences between these categories. In our study, the structure decision point was always on the second noun. We analysed two time-windows in this location, one from 300-500 ms and one from 500-700 ms, to target potential differences in the N400 and P600 which have been previously shown to be sensitive to predictability.

First we discuss the results uncovered in the N400 window. Here we found a three-way interaction of hemisphere, prime type and prime bias match. Follow-up analyses showed that the interaction of prime and bias match was only significant in the left side electrodes and the effect of prime bias match was stronger in the DOD, as opposed to the PD primes. Unpredictable sentences showed a larger N400 in the DOD sentences replicating the N400's predictability-based modulation in our dataset as well (see also Kutas & Hillyard, 1983), but this effect was not significant in the PD sentences.

The N400 results for the DOD primes are in line with an account where participants are continuously predicting the next word while listening to the prime sentences and a processing difference arises if their expectations are not met. Thus, they are compatible with error-based learning theories. These results are also in line with the behavioural results of the current study. It seems that participants' immediately subsequent dative production as well as their neural correlates were sensitive to whether the DOD structure was predictable or surprising. This sensitivity manifested in a bigger N400 effect during the structure decision point of unexpected DOD sentences and also a higher likelihood of structure repetition after unexpected DOD primes.

Some of our EEG results, however, are not in line with the predictions of error based learning theories. These theories predict processing differences in PD sentences as well as DOD sentences. We found no significant ERP differences between the processing of predictable and surprising PD sentences in our dataset, even though (similarly to DODs) participants were more likely to repeat PDs after surprising as opposed to predictable sentences. Decreased sensitivity to the predictability of the more frequent structure has also been found in previous prime surprisal studies. Prime surprisal effects are often weaker (e.g. Peter et al., 2015, current study) or nor detectable (e.g. Jaeger & Snider, 2013) after the more frequent dative structure. Therefore it is also possible that the predictability-related ERP differences in the PD sentences simply did not reach a detectable level in the current study.

Next, we turn to the P600 window. Larger P600 effects typically signal additional syntactic processing (Kaan & Swaab, 2003) or thematic role reassignment (Kuperberg et al., 2003). They tend to appear if the incoming stimuli does not match the previously constructed syntactic structure or previously assigned thematic roles. Thus if participants generate (structure or role) expectations based on the verb's bias, we would expect larger P600 effects at the structure decision point when these expectations are violated in the mis-matching sentences in our study. However, we found no significant differences in this time window. While the N400 effects suggest that participants were sensitive to the predictability of the nouns (in the DOD sentences), we found no evidence of additional syntactic processing, or thematic role re-assignment (typically signalled by a larger P600 effect) in mis-matching sentences. Thus, although the N400 pattern of results is compatible with an account where participants are constantly integrating upcoming words, and where the higher joint verb-structure frequencies (that can also make the words more predictable) led to less effortful integration, the absence of P600 effects means that this study does not support the hypothesis that participants were actively pre-activating the different dative structures or thematic role mappings and that processing costs increased when these expectations were not met. In other words, the lack of P600 effects does not support the possibility that participants made strong syntactic or thematic predictions that were then re-assessed when the incoming stimuli did not match these predictions.

However, before we draw strong conclusions, we need to consider the restrictions of the specific design used. Due to the exploratory nature of the current paradigm (combination of an intervention-based prime surprisal study and EEG), it was harder to accurately estimate the required participant number. Furthermore, as the study's focus was long-term learning, we used a partially between participant design, where each participant only heard sentences from two out of the four prime surprisal conditions (DOD bias group: surprising sentences: PD-biased verb-DOD-structure, predictable sentences: PD-biased verb-PD-structure; PD bias group: surprising sentences: DOD-biased verb-PD-structure, predictable sentences: DOD-biased verb-DOD-structure; see Table 3.2.). This design led to only two (instead of four) per-participant averages per electrode site in each condition for the ANOVAs, resulting in analyses with lower statistical power. Due to the above restrictions replicating the above results and conducting further studies assessing the above questions with larger sample sizes and potentially a fully within-participant design is crucial.

In addition, although this method gives us real-time information on the ERP differences between predictable and surprising sentences, it is not clear what underlying processing mechanisms cause these differences. For instance, while N400 and P600-differences tend to point to differences in predictability, it is debated whether they are the result of some of the words being predicted and some not, or whether they reflect the ease of integration (Kutas & Federmeier, 2011). As prediction (and not integration) related differences are a crucial part of error-based learning theories, it is necessary to differentiate between the two effects in this context. Combining a prime surprisal study with an ERP paradigm measuring differences on words where processing differences can only arise if participants are also anticipating the next word (e.g. DeLong Urbach & Kutas, 2005) could provide more accurate information on whether the prime surprisal effects are connected to predictions (and not only a result of integration differences). For instance, ERP differences could be measured on sentences such as sentences 5.a. and 5.b. where processing differences between *green* and *blue* should only arise if participants both used the DOD-biased verb to anticipate a DOD structure and also took into account previously seen picture stimuli (similarly to the current study) to anticipate the next word of the structure (*cat* versus *dog*), which is predictable in sentence 5.a. but surprising in sentence 5.b. Without predicting the next word, we would not expect processing differences between *green* and *blue* in these sentences. We could then measure whether these potential prediction-related ERP differences are also connected to differences in the likelihood of a priming effect.

5.a. Lisa gave the **green** cat to the blue elephant.

5.b. Lisa gave the **blue** elephant to the green cat.

4.5. Summary and conclusions

To our knowledge the current study is the first study that jointly monitors both the behavioural and neural effects of processing predictable and surprising sentences. It combined a three-stage prime surprisal-based intervention study with ERP measures to monitor both processing differences between predictable and surprising sentences and to assess subsequent sentence production both immediately and in the longer term. Our goal was to evaluate whether the differences between the neural correlates associated with predictable and surprising sentences are compatible with the predictions of error-based theories of language acquisition. Our behavioural results replicated prime surprisal effects found in previous studies showing an increased likelihood of structure repetition after surprising compared to predictable sentences. We also replicated previous results showing that surprising input leads to more learning than predictable input: participants were more likely to produce the dative alternative in the post-test phase that they heard in surprising sentences in the bias phase. Some of our ERP results were also in line with the predictions of error-based learning theories, as we found that surprising DOD sentences elicited a larger N400 effect at the structure decision point than predictable sentences. However, contrary to the predictions of error-based learning theories, we found only much weaker, non-significant, differences in the PD sentences. We also found no P600 differences (typically signalling additional syntactic processing) in the prime sentences where the dative structure did not match the preceding verb's bias. Further examination is required to determine whether P600 effects and N400 differences in the PD sentences could be detected in a study with a fully between-participants design and/or more statistical power.

CHAPTER FIVE: GENERAL DISCUSSION

5.1. Introduction

The goal of this work was to examine the role of prediction in language acquisition. We assessed the workings of linguistic prediction in different linguistic environments and evaluated the projections of the Dual-path model, a connectionist model that operationalises prediction-based learning in syntactic acquisition. We used the prime surprisal paradigm as the basis of our investigations. This method compares structure repetition (priming) after surprising and predictable sentences and is informative about the central claim of error-based learning theories that surprising input is more likely to lead to structure repetition than predictable input.

This thesis contains three experimental chapters. The studies in Chapter 2 examined how versatile linguistic predictions are by assessing whether prime surprisal effects appear with passive as well as dative structures. The main goal of Chapter 3 was to assess a central claim of error-based learning theories that unpredictable input leads to faster learning rates than predictable input. Chapter 4 targeted the neural correlates associated with processing surprising and predictable sentences and examined whether the processing differences between the two sentence types are in line with those suggested by error-based learning theories.

We will begin this chapter by summarising the main findings from the three experimental chapters, then discuss how these results speak to the role of prediction in language acquisition by reviewing whether they support the Dual-path model. We will continue the chapter by discussing the limitations of the present work and conclude the chapter by considering future research directions.

5.2. Summary of findings

The first experimental chapter, Chapter 2, investigated whether prime surprisal appears with syntactic structures other than datives. Previous studies have shown that both adult (e.g. Jaeger & Snider, 2013) and child (Peter et al., 2015) participants are more likely to repeat a dative structure when it is unexpected as opposed to expected, based on the verb's bias. The results of these studies are compatible with the projections of error-based learning theories and support prediction's role in language acquisition as they are in line with the possibility that participants are constantly predicting the next word. On this model, the error-signal resulting from a potential discrepancy between the prediction and the actual input leads to the enhanced likelihood of structure repetition. However, we can only draw limited conclusions based on previous studies as, to our knowledge, published prime surprisal studies to date only featured dative structures. Datives provide a useful test case for assessing error-based learning due to the strong verb biases and the verb's location before the structure decision point. However, other structures with different linguistic characteristics might also be sensitive to prime surprisal. To assess this possibility, we carried out two prime surprisal studies with adult participants. The goal of the first study was to replicate prime surprisal effects with dative structures while the second study aimed to find out whether prime surprisal also appears with sentences involving the voice alternation.

In our dative study we found a numerical priming effect that was comparable to priming effects typically found in the literature (3.9%, see Mahowald et al., 2016) however this effect did not reach significance in our sample. We also found numerical, but non-significant differences in the strength of priming after surprising as opposed to predictable sentences: participants were more likely to produce DODs after mis-matching (PD-biased verb – DOD structure) compared to matching (DOD-biased verb – DOD structure) primes. Prime surprisal, however, was only present in the target sentences following DOD primes, not in those following PD primes.

Unlike in the dative study, we did find a significant priming effect in the voice alternation study. However, we did not find a prime surprisal effect. In this case, unlike in the dative study, there was not even a numerical effect. In other words, we found no evidence of a prime surprisal effect in this study.

As the main focus of these studies was linguistic prediction, we carried out three additional analyses to exclude non prediction-related reasons for the lack of prime surprisal in the voice alternation study. These analyses confirmed that both the prime and target sentences and the verb-bias measures used in the voice alternation study were suitable for demonstrating prime surprisal effects. The absence of prime surprisal in the passive study may then be due to the differences between the dative and transitive structures (in dative structures the verb always proceeds the structure decision point, whereas for transitive structures the order is reversed) or the way the two studies were set up (in the dative study the two priming conditions contained both PD- and DOD-biased verbs, whereas in the passive study all verbs were overall active biased but this bias was weaker in one of the conditions). However, as the prime surprisal method only takes measurements at the beginning (demonstration of surprising or predictable sentence) and endpoints (potential structure repetition) of the learning event, further examination is needed to determine the exact reasons behind the lack of prime surprisal with passive sentences.

The primary goal of Chapter 3 was to target a central, but as yet not directly tested prediction of error-based learning theories that surprising input leads to more learning than predictable input. First, we presented a pilot study testing priming and prime surprisal in dative structures 5-6 year old children, which reported similar results to our prime surprisal study with adult participants described in Chapter 2. We found a significant structural priming effect but no significant prime surprisal effect in either the full dataset or the subsets involving either only DOD or only PD primes. Despite the lack of significant effects the priming effects were numerically larger in the surprising as opposed to predictable sentences. This effect was driven by the difference in the target sentences following DOD primes where participants were 4.92% more likely to produce DODs after surprising as opposed to predictable primes. There was no such difference in the target sentences after PD primes.

We used the results of this pilot study to design the main study presented in Chapter 3. For the main study, we embedded the prime surprisal paradigm in a four-stage intervention study, where we split participants into two bias groups so we could directly compare learning rates after predictable versus surprising input. Participants in the DOD-bias group always heard DODs in surprising and PDs in predictable sentence contexts, while participants in the PD-bias group heard sentences following the opposite pattern. In addition to the main focus of the study (abstract learning rates associated with surprising versus predictable input) we also aimed to replicate immediate priming and prime surprisal effects seen in previous studies. Furthermore, we assessed whether we can also detect verb-specific error-based learning effects.

Our key manipulation of interest (DOD- versus PD-bias group) did not lead to significant differences in the main analyses. However there was a numerical difference in both adult and child groups: participants produced more DODs in DOD- and PDs in the PD-bias group in their post-test compared to their pre-test performance. In addition, this result was significant when we only included participants who did not show ceiling performance in the pre-test phase, and for the child group alone when age and syntactic knowledge (measured by the TROG test) was taken into account. In essence, these learning effects were significantly stronger for younger children.

Our results did not show the expected patterns in the other phases of the study. While in the bias phase adult participants were more likely to repeat the previous dative structure when it was unexpected (if their baseline DOD performance was taken into account), children showed no immediate prime surprisal effects. Our exploratory analyses confirmed that the lack of prime surprisal effect in the child group was not due to the overall DOD- or PD-biased bias-phase, the increasing predictability of verb-structure pairings, or the variable baseline DOD-production rates. Another phase that did not lead to the expected pattern, was the second post-test. Here all participants were more likely to repeat the dative structure that verbs appeared with earlier, but the likelihood of structure repetition did not depend on whether the structure was surprising or predictable in the bias phase. Thus, neither the child nor the adult group showed verb-dependent error-based learning effects. However, we cannot draw strong conclusions based on the absence of these effects in the current work, as this phase of the study was exploratory and provided a less strong test of learning than the main test of abstract learning. Overall the results supported the main research question (albeit only in exploratory analyses) targeting learning rates for abstract structures after predictable and surprising input, but was inconclusive with respect to the secondary questions. This might reflect the nature of the design, in that we prioritised the main question and this led to compromises in the assessment of the secondary questions.

The goal of Chapter 4 was to gather more on-line information about the processing of predicable and surprising dative sentences similar to the ones that lead to immediate prime surprisal and learning effects in the previous studies. We conducted a three-stage prime surprisal-based intervention study (modeled after the main study reported in Chapter 3) where we monitored adult participants' ERPs during sentence processing and also compared both their immediate and delayed sentence production after more or less predictable sentences.

The behavioural results were similar to those found in previous studies: participants were more likely to produce DODs after DOD (as opposed to PD) primes and this priming effect was larger after surprising as opposed to predictable sentences, demonstrating a prime surprisal effect. We found no significant differences in participants' post-test performance depending on bias group. However, we found numerical differences similar to those in Chapter 3, as participants were more likely to produce the dative structure in the post-test phase that they were exposed to with mis-matching verbs in the bias-phase. It is possible that the long-term learning effects did not yield significant results in this study because as the main target of the study was the EEG assessment, it did not have enough statistical power for this comparison.

The EEG analyses reported in Chapter 4 targeted the sentence region where error-based learning mechanisms predict processing differences: the structure decision point. Here it becomes clear to the participants whether they are hearing a DOD or PD structure and they can also assess whether their verb-based structure predictions were met or not. We assessed the regions 300-500 and 500-700 ms after word onset to assess potential N400 and P600 differences. In the N400 window we found the three-way interaction of hemisphere, prime type and prime bias match. Follow-up analyses showed that the interaction of prime and bias match was only significant in the left side electrodes and the effect of prime bias match was stronger in the DOD, as opposed to the PD primes. Unpredictable (mis-matching) sentences showed a larger N400 in the DOD sentences, but this effect did not appear in the PD sentences. We found no P600 differences in the 500-700 ms post-word onset window. While the N400 differences suggest that participants were sensitive to the predictability of the nouns (in the DOD sentences) this study does not provide any evidence for additional processing due to syntactic re-analyses or thematic role re-assignment (typically signalled by a larger P600 effect) in mis-matching sentences. In the next section we will discuss how these results relate to the projections of error-based theories of language acquisition.

5.3. Discussion of the findings in light of error-based learning theories

Next we will discuss our findings in light of prediction's role in language acquisition. In particular we will consider whether our results are in line with the projections of the Dual-path model (Chang, Dell & Bock, 2006), the model that operationalises syntactic acquisition via prediction-based learning. This connectionist, frequency-based model suggests that both children and adults continuously predict upcoming words when listening to others talk. They then compare their predictions to the actual input, and if there is a discrepancy, an error signal is generated, which is then used to adjust the weights supporting syntactic knowledge. These weight changes lead to an increased likelihood of immediate structure repetition and they also accumulate over time. This accumulation allows children's syntactic knowledge to gradually approximate the adult state and adults' knowledge to adapt to different linguistic situations or registers.

Prime surprisal, the method we used as the starting point of the current investigations, has already provided support for the Dual-path model. Previous studies showed that both adults and children are sensitive to immediate prime surprisal (e.g. Jaeger & Snider, 2013; Peter et al., 2015). Some of our studies comparing priming after surprising and predictable dative sentences have also replicated immediate prime surprisal effects with adult participants. We found some evidence of prime surprisal throughout the studies, providing further support for the Dual-path model. However, prime surprisal effects did not appear in all parts of our investigations.

In Chapter 2, the Dual-path model should predict prime surprisal effects in both studies involving the dative and the voice alternation. But while datives showed numerical, but non-significant prime surprisal in our first study, passive sentences in the second study did not show such effects. Based on these results alone we cannot determine what exactly prevented the appearance of prime surprisal in the voice alternation study, but this pattern results is not in line with the predictions of the Dual-path model.

Furthermore, while the 5-6 year old child participants showed enhanced learning after surprising as opposed to predictable stimuli in Chapter 3 in our exploratory analyses, they did not show any evidence of immediate prime surprisal effects. As Peter and colleagues (2015) previously found prime surprisal with the same age-group using a similar design, it is possible that the specific learning-focused modifications in our study interfered with the immediate prime surprisal effects. Nevertheless, these results are challenging for the Dual-path theory which posits that immediate prime surprisal effects and long-term learning effects are the result of the same error-based learning mechanism. The dissociation of these effects either suggests that they may not always be caused by the same mechanism or requires further explanation from this theory.

The longer-term learning effects themselves, reported in Chapter 3, however, provided important preliminary evidence in favour of the Dual-path model, however, these differences did not reach significance in our pre-registered analyses. We found that participants in both the child and the adult group produced numerically more DODs in DOD- and PDs in the PD-bias group compared to their pre-test performance showing that surprising input leads to faster learning rates than predictable input. To our knowledge this was the first study directly comparing learning rates associated with surprising versus predictable input, confirming that predictability not only has immediate but also long-term effects on linguistic behaviour. We also found a similar pattern of results in Study 3, where adult participants were numerically more likely to produce DODs in the DOD as opposed to the PD bias group (however, this result did not reach significance in this study either), providing further preliminary evidence for predictability's role in language learning.

Chapter 3 also provided information about another prediction of the Dual-path theory, that learning rates decrease with age. According to this model as learners accumulate more knowledge, their syntactic representations become stronger and thus less sensitive to interventions. Our results were consistent with this prediction in two ways. First, the pre-to post-test shift was numerically larger in the child versus the adult group, although this effect did not reach significance. Furthermore, within the child group, younger children were more sensitive to our bias manipulation indicating an enhanced learning rate.

This pattern however did not appear with the immediate prime surprisal effects. Prime surprisal was larger in the adult group than in the child group (where it was not detectable) and these effects were not dependent on age within the child group either. As, according to the Dual-path model, immediate prime surprisal is the result of the same error-based learning mechanism that leads to long-term learning, this pattern of results is not in line with the predictions of this model.

The main study in Chapter 3 uncovered another result that could be challenging for the Dual path model: we found no verb-based learning effects in the main study in Chapter 3 in either the adult or the child group. It is however important to note, that both the immediate prime surprisal and the verb-based learning effects may have been influenced by our between-participant design and thus are less conclusive than those concerning abstract learning. It is thus crucial to target these effects in a study where they are the primary focus of the investigation.

Finally, we turn to the results of our EEG study. This study replicated immediate prime surprisal effects and enhanced learning after surprising versus predictable input with adults (thus replicating the pattern of results found in Chapter 3 in support of the Dual-path model). Crucially we also found enhanced N400 effects in surprising versus predictable DOD (but not PD) sentences but found no detectable differences in the P600 window. While the N400 effects suggest that participants were sensitive to the predictability of the dative structure (in the DOD sentences), we found no evidence of additional syntactic processing, or thematic role re-assignment (typically signalled by a larger P600 effect) in mis-matching sentences. This pattern of results is compatible with an account where participants are constantly integrating upcoming words and the higher joint verb-structure frequencies (that can also make the words more predictable) lead to less effortful integration evidenced by the larger N400 effect. However, without P600 differences this study does not support the possibility that participants were actively pre-activating the different dative structures or thematic role mappings and processing costs increased when these expectations were not met. In other words, in the absence of P600 effects we found no support for the possibility that participants made strong syntactic or thematic predictions that were then re-assessed when the incoming stimuli did not match these predictions.

Overall this work granted some support for predictability's role in language learning by providing preliminary evidence that less predictable sentences lead to an enhanced learning rate compared to predictable ones and it also replicated effects showing that this is also the case in the short term with adults. However, as the above results only reached significance in exploratory analyses they need to be replicated with appropriate sample sizes. The remaining results are less clearly in favour of error-based theories. We did not show immediate prime surprisal with child participants or with structures other than datives. We also did not find that verb-based learning effects were sensitive to predictability. While some of these behavioural results are less conclusive as they were not the main focus of the studies in question, they still are challenging for the Dual-path theory and require further examination. Our ERP study added further details to the picture as it showed processing differences related to predictability at the structure decision point of the dative sentences (where the Dual-path model would predict such differences), but they are inconclusive as to whether these processing differences are the result of active predictions or rather differences in the ease of integration. Overall, these results suggest that while behavioural and neuropsychological patterns predicted by error-based theories can be detected in various situations, the mechanisms leading to these patterns may not always operate in the ways described by the Dual-path model.

5.4. Limitations

This work has two overarching limitations that are relevant to all of our studies: the first comes from the exploratory nature of the current work, and the second is the result of limitations of the paradigm (prime surprisal) we used as our starting point. First, as two out of three chapters contained studies that combined the prime surprisal method with other paradigms they have not been used with before (intervention study, ERP study) it was especially challenging to determine both the appropriate exclusion criteria (see participants with a pre-test ceiling performance in Chapter 3) and participant numbers (see Chapters 3 and 4). As a result of this, some of the results which showed informative patterns only reached significance in exploratory analyses or not at all. In these cases it cannot be determined based on the current work alone whether these results did not reach significance because they are unreliable or because the study lacked adequate statistical power. It is crucial to replicate these results in future work. The overall exploratory nature of this work has also led to some limitations that are specific to the main study reported in Chapter 3. While the decision to prioritize exploring the main abstract learning effect was essential to ensure that the potential differences is our main comparisons were due to different levels of predictability, it also meant that the results of our secondary comparisons (immediate prime surprisal and verb-based learning effects) led to less conclusive results.

Another important limitation of the current work lies in the nature of the method that we used as our starting point, the prime surprisal paradigm. This method is especially useful to target error-based learning as it can directly address the potential changes in language production depending on the predictability of the input. But as it only takes measurements at the beginning and the end of the learning event (when the predictable versus surprising input is presented and when the subsequent production occurs), it does not provide us any information about how or why surprising sentences might lead to more language change. It is therefore crucial to pair this method with on-line measures such as EEG or eye-tracking to gain more information about the details of how these learning effects unfold over time. We aimed to address this last limitation by embedding the intervention study described in Chapter in an EEG study in Chapter 4. While this study provided important additional information on the processing of surprising and predictable sentences, EEG itself also has its own limitations when it comes to assessing predictions. Although it gives us real-time information on the ERP differences between predictable and surprising sentences, it is not clear what underlying processing mechanisms cause these differences. For instance, while the N400 differences found in DOD sentences in the study reported in Chapter 4 are typically in sync with differences in predictability, it is debatable whether they are the results of some of the words being predicted and some not, or reflect the ease of integration (Kutas & Federmeier, 2011). As prediction- (and not integration-) related differences are a crucial part of error-based learning theories, future research is necessary to differentiate between the two effects in this context.

While the methodological limitations resulting from using novel methods made it difficult to draw strong conclusions based on some of our results, this work has helped establish a paradigm that can be used to target error-based learning in children in future work. This paradigm can be adapted to target a range of different aspects of error-based learning theories of language acquisition. We will discuss some of these in our last section, considering future research directions.

5.5. Directions for future research

While this work has uncovered important information about prediction's role in language acquisition, it naturally also raised several additional research questions. First, the aim of Chapter 2 (prime surprisal with the dative and the voice alternation) was to examine under what circumstances prime surprisal appears. While we found numerical, but non-significant differences between priming after surprising and predictable datives sentences, sentences involving the voice alternation showed no sign of a prime surprisal effect. However, as there are several differences between our passive prime surprisal study and other studies demonstrating prime surprisal effects with datives, future research is necessary to determine which processing differences led to the different results. Conducting further prime surprisal studies using different structures and set-ups would also provide additional information about the circumstances in which linguistic predictions occur and about how they might contribute to learning. For instance, we could assess whether sentences involving the voice alternation lead to prime surprisal if we include overall active and overall passive-biased verbs (instead of more and less active biased verbs as in the current study). We could also examine whether prime surprisal appears with passive sentences if we do not use pictures to disambiguate between the agent and the patient, thus changing the location of the structure decision point. It may also be possible to find priming differences between sentences that are overall either PD- or DOD-biased, but to different extents. Answering these questions will tell us more about how the location of the verb and structure decision point influences prime surprisal and also about whether people are only sensitive to overall biases or to the extent of these biases as well. Another advantage of finding prime surprisal effect with structures other than datives would be that we could then also assess if these structures also lead to both immediate and cumulative, longer-term structure repetition, such as datives did in Chapter 3. In order to show that error-based learning is a viable and widespread learning mechanism, it is crucial to demonstrate that it operates with different structures (not just datives) and if it does not affect some structures it is important to explain why.

While the main study in Chapter 3 provided some preliminary evidence for prediction's role in language acquisition and for error-based learning theories, we can only draw limited conclusions based on one study. First, the main results of this study (enhanced learning rates with surprising as opposed to predictable stimuli) need to be replicated in studies where the power calculations determining the sample sizes take into account the restrictions determined by the current study. Furthermore, the scope of these effects also needs to be assessed in order to gain a more complete picture of these mechanisms. For instance, it is essential to assess whether children younger than those featured in our study are also sensitive to these effects, how permanent these effects are and what structures they appear with. Additionally, since we found no significant relationship between the learning effects found in Chapter 3 and syntactic knowledge measured by the TROG test, it is also essential to assess whether other tests of syntactic knowledge or other potentially relevant aspects of linguistic and cognitive abilities (e.g. vocabulary size or working memory) correlate with these effects to see what factors might contribute to error-based learning.

As noted in the previous section, a specific limitation of this work was that while it assessed multiple effects that are relevant to error-based learning, the results concerning immediate prime surprisal and verb-based learning are less conclusive than those concerning abstract error-based learning. Thus, additional studies will need to assess whether verb-based learning is also sensitive to structure predictability and whether the dissociation of the immediate prime surprisal and longer-term learning effects in the child group was only due to the specific design used in the current work. Carrying out further work targeting the relationship of immediate priming and long-term learning is also crucial to better understand the mechanisms causing priming itself. This has been the topic of previous debate (see e.g. Pickering & Branigan; 1998 and Chang, Janciauskas, & Fitz, 2012) as determining the mechanisms driving immediate structure repetition is also crucial for understanding language processing itself, not just language acquisition.

While this work has provided important preliminary evidence showing that less predictable input leads to enhanced learning rates, it has also raised questions about the precise nature of the mechanism that leads to this effect. For instance, the absence of the prime surprisal effects in Chapter 2 raises doubts about whether error-based learning is only sensitive to overall verb biases or to the relative strengths of these biases as well. Furthermore, the child results in Chapter 3 question whether immediate priming and longer-term learning effects are always caused by the same mechanism. However, perhaps the most important question is the one highlighted by Chapter 4 providing on-line information about the processing of predictable and surprising dative sentences. While this study showed predictability-related processing differences at the structure decision point of dative sentences, these studies alone cannot determine whether these differences were elicited by active predictions or not. As predictions are a crucial part of error-based learning theories, future research will have to determine whether active predictions are a part of the mechanism leading to these effects.

5.6. Concluding remarks

Overall this work examined prediction's role in language acquisition by expanding on results obtained by the prime surprisal paradigm and by designing novel studies building on this method. We examined whether prime surprisal appears with passive as well as dative structures in adults, whether surprising sentences lead to enhanced learning rates and finally by comparing the ERPs associated with predictable and surprising sentences to assess whether the potential processing differences are in line with those projected by error-based theories of language acquisition.

Some of our key results support error-based learning theories. We found preliminary evidence that less predictable input not only influences subsequent language production in the short term (a result that we also replicated in all of our studies with adult participants), but that it also leads to longer-term, cumulative effects. The above effects need to be interpreted with caution as they only reached significance in our exploratory analyses. However, if replicable, they will provide the first evidence that less predictable input leads to more learning than predictable input, providing crucial evidence for prediction's role in language acquisition and in support of error-based theories of language acquisition.

Some of our results are not in line with the specific projections of error-based models. For instance, while the learning rates associated with abstract structures were sensitive to predictability, the rate of verb-dependent learning was not. Furthermore, while we replicated dative-based prime surprisal effects with adults in some of our studies, we did not find such effects with either passives structures or with child participants. While some of these results might be influenced by the limitations of the current studies, the specific reasons behind the pattern of results in our behavioural studies require further examination. Our ERP results also provided inconclusive evidence for error-based learning theories. While they showed predictability-related processing differences at the structure decision point of the dative sentences, these results alone cannot determine whether these differences are the result of active prediction or differences in the ease of integration. Future research will have to examine the exact circumstances in which immediate prime surprisal and predictability-dependent learning effects appear and, crucially, uncover the exact nature of the learning mechanism leading to these effects. As well as widening our knowledge about prediction's role in language acquisition, an important contribution of the present work was in establishing an experimental paradigm that can directly target different aspects of error-based theories of language acquisition. This paradigm could form the basis for future research in this topic.

BIBLIOGRAPHY

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition, 73(3), 247-264.

Ambridge, B., Pine, J. M, Rowland, C. F., Freudenthal, D., Chang, F. (2014). Avoiding dative overgeneralization errors: semantics, statistics or both? Language, Cognition and Neuroscience, 29(2), 218-243

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language, 57(4), 502-518.

Baayen, R. H., Davidson, D. J., Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390–412.

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. Neuroreport, 15(3), 545-548.

Barr, D. J., Levy, R., Scheepers, C., Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68, 255–278. (doi:10.1016/j.jml.2012.11.001)

Bencini, G. M. L. & Valian, V. V. (2008). Abstract sentence representations in 3-year olds: Evidence from language production and comprehension. Journal of Memory and Language, 59, 97-113.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? Cognition, 114(3), 455-461.

Bock, K. (1986). Syntactic persistence in language production. Cognitive Psychology, 18, 355-387.

Bock, K. (1989). Closed-class immanence in sentence production. Cognition, 31, 163-186.

Bock, K. & Loebell, H. (1990). Framing sentences. Cognition, 35, 1-39.

Bock, K., Dell, G. S., Chang, F., & Onishi, K. H. (2007). Persistent structural priming from language comprehension to language production. Cognition, 104(3), 437-458.

Borovsky, A., Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. Developmental Psychology, 50, 1600–1613.

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. Journal of Experimental Child Psychology, 112(4), 417-436.

Borovsky, A., Sweeney, K., Elman, J. L., & Fernald, A. (2014). Real-time interpretation of novel events across childhood. Journal of Memory and Language, 73, 1-14.

Branigan, H. P., Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. Cognition, 157, 250-256.

Branigan, H. P., Pickering, M. J, & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. Psychonomic Bulletin & Review, 6, 635–640.

Branigan, H. P., Pickering, M. J, & McLean, J. F. (2005). Priming prepositional-phrase attachment during comprehension. Journal of Experimental Psychology: Learning, Memory, & Cognition, 31, 468-481.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In Cognitive Foundations of Interpretation. KNAW, 69-94.

Brothers, T., Dave, S., Hoversten, L. J., Traxler, M. J., & Swaab, T. Y. (2019). Flexible predictions during listening comprehension: Speaker reliability affects anticipatory processes. Neuropsychologia, 135, 107225.

Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. Applied Psycholinguistics, 34(3), 519-539.

Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. Journal of Memory and Language, 66(4), 833-849.

Campbell, A. L. Tomasello, M. (2001). The acquisition of English dative constructions. Applied Psycholinguistics, 22: 253–267.

Casillas, M., Bobb, S. C., Clark, E. V. (2016). Turn taking, timing, and planning in early language acquisition. Journal of Child Language, 43, 1310–1337.

Chang, F., Dell, G. S., Bock, K. (2006). Becoming syntactic. Psychological review, 113, 234. (doi:10.1037/0033-295X.89.1.1)

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. Language and Linguistics Compass, 6/5, 259-278.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36, 181-204.

Christiansen, M. H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. Cognitive Science, 23(2), 157-205.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. Behavioral and Brain Sciences, 39.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37–46.

Conwell, E., Demuth, K. (2007). Early syntactic productivity: Evidence from dative shift. Cognition, 103, 163-179.

De Ruiter, J. P., Mitterer, H., Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. Language, 82, 515–535.

Dell, G. S., Chang, F. (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. Phil. Trans. R. Soc. B, 369, 20120394. (doi:10.1098/rstb.2012.0394)

Delaney-Busch, N., Morgan, E., Lau, E., & Kuperberg, G. R. (2019). Neural evidence for Bayesian trial-by-trial adaptation on the N400 during semantic priming. Cognition, 187, 10-20.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nature Neuroscience, 8(8), 1117.

Elman, J. L. (1990). Finding structure in time. Cognitive science, 14(2), 179-211.

Fine, A. B., & Jaeger, T.F. (2013). Evidence for implicit learning in syntactic comprehension. Cognitive Science, 37(3), 578-591.

Fine, A. B., Jaeger, T. F., Farmer, T. A., Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. PloS one, 8, e77661.

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. Brain and language, 140, 1-11.

Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. Cognition, 157, 340-351.

GNU Image Manipulation Program, The GIMP team, GIMP 2.8.10, www.gimp.org, 1997-2014.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. Journal of Phonetics, 38(4), 555-568.

Hilbrink, E. E., Gattis, M., Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: A longitudinal study of mother–infant interaction. Frontiers in Psychology, 6, Article 1492. doi:10.3389/fpsyg.2015.01492.

Hirsh-Pasek, K. & Golinkoff, R. (1996). The origins of grammar. Cambridge, MA: MIT Press.

Hoedemaker, R. S., Meyer, A. S. (2018). Planning and coordination of utterances in a joint naming task. Journal of Experimental Psychology: Learning, Memory, and Cognition. Advance online publication.

Hohenstein, S., & Kliegl, R. (2013). Remef (REMove effects) (version v0.6.10). <http://read.psych.uni-potsdam.de/joomla/attachments/article/12/ remef.v0.6.10.R>.

Holcomb, P. J., Grainger, J., & O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. Journal of Cognitive Neuroscience, 14(6), 938-950.

Huang, H. W., Meyer, A. M., & Federmeier, K. D. (2012). A “concrete view” of aging: Event related potentials reveal age-related changes in basic integrative processes in language. Neuropsychologia, 50(1), 26-35.

Huang, Y. T., Zheng, X., Meng, X., Snedeker, J. (2013). Children’s assignment of grammatical roles in the online processing of Mandarin passive sentences. Journal of Memory and Language, 69. (doi:10.1016/j.jml.2013.08.002)

Huettig, F., & Guerra, E. (2015). Testing the limits of prediction in language processing: Prediction occurs but far from always. In the 21st Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP 2015).

Huettig, F., Mani, N. (2016). Is prediction necessary to understand language? Probably not. Language, Cognition and Neuroscience, 31(1), 19-31.

Indefrey, P., Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. Cognition, 101–144.

Ito, A., Corley, M., & Pickering, M. J. (2018). A cognitive load delays predictive eye movements similarly during L1 and L2 comprehension. Bilingualism: Language and Cognition, 21(2), 251-264.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language, 59, 434–446.

Jaeger, T. F., Snider, N. E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime’s prediction error given both prior and recent experience. Cognition, 127, 57–83.

Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In Proceedings of the Cognitive Science Society Conference (pp. 1061-1066).

Kamide, Y., Altmann, G. T., & Haywood, S. L. (2005). 10 The Time Course of Constraint Application during Sentence Processing in Visual Contexts: Anticipatory Eye Movements in English and Japanese. Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions, 229.

Kantola, L., & van Gompel, R. P. (2011). Between-and within-language priming is the same: Evidence for shared bilingual syntactic representations. Memory & Cognition, 39(2), 276-290.

Kaschak, M. P., Kutta, T. J., Jones, J. L. (2011). Structural priming as implicit learning: Cumulative priming effects and individual differences. Psychonomic bulletin & review, 18, 1133-1139.

Kidd, E. (2012). Individual differences in syntactic priming in language acquisition. Applied Psycholinguistics, 33, 393-418.

Kukona, A., Fang, S. Y., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. Cognition, 119(1), 23-42.

Kukona, A., Cho, P. W., Magnuson, J. S., & Tabor, W. (2014). Lexical interference effects in sentence processing: Evidence from the visual world paradigm and self-organizing models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(2), 326.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension?. Language, cognition and neuroscience, 31(1), 32-59.

Kutas, M., Federmeier, K. D. (2011) Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annual review of psychology, 62, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. Memory & cognition, 11(5), 539-550.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307(5947), 161.

Levy, R. (2008). Expectation-based syntactic comprehension. Cognition, 106(3), 1126-1177.

Levinson, S. C. (2016). Turn-taking in human communication–origins and implications for language processing. Trends in Cognitive Sciences, 20(1), 6-14.

Lew-Williams, C., Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. Psychological Science, 18, 193–198.

Lin, Y., Fisher, C. (2017). Error-Based Learning: A Mechanism for Linking Verbs to Syntax. Paper presented at the 39th Annual Meeting of the Cognitive Science Society.

Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers Optimize the Timing of Their Conversational Turns Through Flexible Coordination of Language Comprehension and Production. Psychological Science, 30(4), 504-515.

Lukyanenko, C., Fisher, C. (2016). Where are the cookies? Two- and three-year olds use number-marked verbs to anticipate upcoming nouns. Cognition, 146, 349-370.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Magyari, L., & de Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. Frontiers in Psychology, 3, 376.

Magyari, L., Bastiaansen, M. C., de Ruiter, J. P., Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. Journal of Cognitive Neuroscience, 26, 2530–2539.

Mahowald, K., James, A., Futrell, R., & Gibson, E. (2016). A meta-analysis of syntactic priming in language production. Journal of Memory and Language, 91, 5-27.

Mani, N., Daum, M. M., Huettig, F. (2016). "Proactive" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. The Quarterly Journal of Experimental Psychology, 69, 2189-2201.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. Journal of Experimental Psychology: Human Perception and Performance, 38(4), 843.

Mani, N., Daum, M. M., & Huettig, F. (2016). "Proactive" in many ways: Developmental evidence for a dynamic pluralistic approach to prediction. Quarterly Journal of Experimental Psychology, 69(11), 2189-2201.

Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., Taylor, A., (1999). Treebank-3.

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. Journal of Memory and Language, 69, 574 – 588.

Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. Journal of Memory and Language, 66, 568-586.

Mitsugi, S., & MacWhinney, B. (2016). The use of case marking for predictive processing in second language Japanese. Bilingualism: Language and Cognition, 19(1), 19-35.

Nation, K., Marshall, C. M., Altmann, G. T. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. Journal of Experimental Child Psychology, 86, 314-329.

Neter, J., Wasserman, W., Kutner, M. H. (1985). Applied linear statistical models: Regression, analysis of variance, and experimental designs. Homewood, IL: Irwin.

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N. & Mézière, D. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. ELife, 7, e33468.

Nieuwland, M., Barr, D., Bartolozzi, F., Busch-Moreno, S., Donaldson, D., Ferguson, H. J., & Ito, A. (2019). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. Proceedings of the Royal Society B: Biological Sciences.

Novembre, G., Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. Consciousness and cognition, 20, 1232-1243.

Olguin, R. & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. Cognitive Development, 8, 245-272.

O'Reilly, R. C., Wyatte, D. R., Rohrlich, (2017). Deep Predictive Learning: A Comprehensive Model of Three Visual Streams. arXiv preprint arXiv:1709.04654.

Peter, M. S. (2015). The Role of the Verb in the Development of Syntax: Evidence from the Structural Priming Paradigm (Doctoral dissertation, University of Liverpool).

Peter, M., Chang, F., Pine, J. M., Blything, R., Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. Journal of Memory and Language, 81, 1-15.

Pickering, M. J. & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. Journal of Memory and Language, 39, 633-651.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. Psychological Bulletin, 144(10), 1002.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. Behavioral and Brain Sciences, 36(4), 329-347.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2000). Available from http://www.pstnet.com.

Rabagliati, H., Gambi, C., Pickering, M. J. (2016). Learning to predict or predicting to learn? Language, Cognition and Neuroscience, 31, 1-12.

Ramscar, M., Dye, M., McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. Language, 89, 760–793.

R Core Team R: A language and environment for statistical computing. 3-900051-07-0. Vienna, Austria: R Foundation for Statistical Computing. 2012 http://www.R-project.org/

Rabagliati, H., Gambi, C., Pickering, M. J. (2016). Learning to predict or predicting to learn? Language, Cognition and Neuroscience, 31, 1-12.

Rumelhart, D. E., Hinton, G. E., Williams, R. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland, editors. Parallel distributed processing. Explorations in the microstructure of cognition. Cambridge: MIT Press. 1. 318– 362.

Rowland, C., Chang, F., Ambridge, B., Pine, J. M., & Lieven, E.V.M. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. Cognition, 125, 49-63.

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of children's linguistic representations: lexical and structural priming of syntactic structures in young children. Developmental Science, 6, 557-567.

Savage, C., Lieven, E., Theakston, A., Tomasello, M. (2006). Structural priming as implicit learning in language acquisition: The persistence of lexical and structural priming in 4- year-olds. Language Learning and Development, 2, 27-49.

Stahl, A. E., Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. Science, 348, 91-94. (doi:10.1126/science.aaa3799)

Stivers, T., Sidnell, J., Bergen, C. (2018). Children's responses to questions in peer interaction: A window into the ontogenesis of interactional competence. Journal of Pragmatics, 124, 14–30.

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. Journal of Memory and Language, 24(2), 232-252.

Thothathiri, M., & Snedeker, J. (2008). Give and take: Syntactic priming during spoken language comprehension. Cognition, 108(1), 51-68.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. Journal of Psycholinguistic Research, 29(6), 581-595.

Traxler, M. J., Tooley, K. M., & Pickering, M. J. (2014). Syntactic priming during sentence comprehension: Evidence for the lexical boost. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(4), 905-918.

Twomey, K., Chang, F., & Ambridge, B. (2014). Do as I say, not as I do: A lexical distributional account of English locative verb class acquisition. Cognitive Psychology, 73, 41-71.

Urgesi, C., Savonitto, M. M., Fabbro, F., Aglioti, S. M. (2012). Long-and short-term plastic modeling of action prediction abilities in volleyball. Psychological research, 76, 542-560.

Wendorf, C. A. (2004). Primer on multiple regression coding: Common forms and the additional case of repeated contrasts. Understanding Statistics, 3, 47–57.

Wlotko, E. W., & Federmeier, K. D. (2012). Age-related changes in the impact of contextual strength on multiple aspects of sentence comprehension. Psychophysiology, 49(6), 770-785.

Yan, S., Kuperberg, G. R., & Jaeger, T. F. (2017). Prediction (or not) during language processing. A commentary on Nieuwland et al.(2017) and DeLong et al.(2005). bioRxiv, 143750.

Yazbec, A., Kaschak, M. P., & Borovsky, A. (2019). Developmental Timescale of Rapid Adaptation to Conflicting Cues in Real-Time Sentence Processing. Cognitive Science, 43(1), e12704.

Ziegler, J., Bencini, G., Goldberg, A., & Snedeker, J. (2019). How abstract is syntax? Evidence from structural priming. PsyRxiv Preprints.

APPENDIX

6.1. Sentence lists for Study 1a and Study 1b

6.1.1. Dative study – Study 1a

	Verb	Bias	Sentences (DOD version)
Prime	Bring	PD biased	The prince brought the princess a baby.
Prime	Bring	PD biased	Wendy brought Bob a bird.
Prime	Send	PD biased	The girl sent the boy a fish.
Prime	Send	PD biased	Dora sent Boots a rabbit.
Prime	Take	PD biased	The king took the queen a cat.
Prime	Take	PD biased	Piglet took Tigger a puppy.
Prime	Sell	PD biased	The student sold the teacher a monkey.
Prime	Sell	PD biased	Marge sold Homer a butterfly.
Prime	Leave	PD biased	The doctor left the nurse a mouse.
Prime	Leave	PD biased	Lisa left Bart a fox.
Prime	Hand	DOD biased	The king handed the queen a cat.
Prime	Hand	DOD biased	Piglet handed Tigger a puppy.
Prime	Offer	DOD biased	The doctor offered the nurse a mouse.
Prime	Offer	DOD biased	Lisa offered Bart a fox.
Prime	Award	DOD biased	The student awarded the teacher a monkey
Prime	Award	DOD biased	Marge awarded Homer a butterfly.
Prime	Give	DOD biased	The girl gave the boy a fish.
Prime	Give	DOD biased	Dora gave Boots the rabbit.
Prime	Show	DOD biased	The prince showed the princess a baby.
Prime	Show	DOD biased	Wendy showed Bob a bird.
Target	Throw	Equibalanced	The king threw the queen a cat.
Target	Throw	Equibalanced	Wendy threw Bob a fox.
Target	Throw	Equibalanced	The girl threw the boy a fish.
Target	Throw	Equibalanced	Dora threw Boots a rabbit.
Target	Drop	Equibalanced	The prince dropped the princess a baby.
Target	Drop	Equibalanced	Piglet dropped Tigger a puppy.
Target	Drop	Equibalanced	The student dropped the teacher a monkey.
Target	Drop	Equibalanced	Marge dropped Homer a butterfly.
Target	Flick	Equibalanced	The king flicked the queen a cat.
Target	Flick	Equibalanced	Wendy flicked Bob a fox.
Target	Flick	Equibalanced	The doctor flicked the nurse a mouse.
Target	Flick	Equibalanced	Dora flicked boots a rabbit.
Target	Slide	Equibalanced	The prince slid the princess a baby.
Target	Slide	Equibalanced	Piglet slid Tigger a puppy.
Target	Slide	Equibalanced	The doctor slid the nurse a mouse.
Target	Slide	Equibalanced	Lisa slid Bart a fox.
Target	Feed	Equibalanced	The girl fed the boy a fish.
Target	Feed	Equibalanced	Lisa fed Bar a fox.
Target	Feed	Equibalanced	The student fed the teacher a monkey.
Target	Feed	Equibalanced	Marge fed Homer a butterly.
Filler			Boots was flying.

Filler	The princess jumped.
Filler	Piglet and Tigger bounced.
Filler	The king and queen waved.
Filler	Tigger was washing.
Filler	The prince was rocking.
Filler	Piglet waved.
Filler	The cat was swinging.
Filler	Dora was flying.
Filler	Bob was swinging.
Filler	The princess and the cat were rocking.
Filler	Dora and Boots waved.
Filler	Bob was flying.
Filler	The prince jumped.
Filler	Tigger was rocking.
Filler	The queen waved.
Filler	The king and queen bounced.
Filler	Piglet jumped.
Filler	Wendy was flying.
Filler	Dora was washing.
Filler	The boy waved.
Filler	Boots pointed at Dora.
Filler	Wendy and Bob jumped.
Filler	Dora was swinging.
Filler	The girl waved.
Filler	Wendy pointed at Bob.
Filler	Boots was washing.
Filler	Piglet was rocking.
Filler	The cat bounced.
Filler	Bob jumped.
Filler	Boots waved at Dora and the baby.
Filler	The king pointed at the queen.
Filler	The doctor and the nurse waved.
Filler	The mouse was rocking.
Filler	The student pointed at the teacher.
Filler	The student was swinging.
Filler	Homer was washing.
Filler	Homer and Marge bounced.
Filler	Lisa and the fox were flying.
Filler	Bart jumped.

6.1.2. Transitive study – Study 1b

	Verb	Bias	Sentences (passive version)
Prime	Disgust	More passive	The teacher was disgusted by the student.
Prime	Disgust	More passive	Lisa was disguested by Bart.
Prime	Fascinate	More passive	The prince was fascinated by the princess.

Prime	Fascinate	More passive	Homer was fascinated by Marge.
Prime	Sadden	More passive	The king was saddened by the queen.
Prime	Sadden	More passive	Wendy was saddened by Bob.
Prime	Aggravat e	More passive	The doctor was aggravated by the nurse.
Prime	Aggravat e	More passive	Piglet was aggaravated by Tigger.
Prime	Dazzle	More passive	The boy was dazzled by the girl.
Prime	Dazzle	More passive	Dora was dazzled by Boots.
Prime	Scare	Less passive	The prince was scared by the princess.
Prime	Scare	Less passive	Lisa was scared by Bart.
Prime	Bother	Less passive	The queen was bothered by the king.
Prime	Bother	Less passive	Wendy was bothered by Bob.
Prime	Surprise	Less passive	The boy was surprised by the girl.
Prime	Surprise	Less passive	Piglet was surprised by Tigger.
Prime	Impress	Less passive	The doctor was impressed by the nurse.
Prime	Impress	Less passive	Dora was impressed by Boots.
Prime	Shock	Less passive	The teacher was teased by the student.
Prime	Shock	Less passive	Homer was teased by Marge.
Target	Irritate	Equibalanced	The teacher was irritated by the student.
Target	Irritate	Equibalanced	Wendy was irritated by Bob.
Target	Irritate	Equibalanced	The prince was irritated by the princess.
Target	Irritate	Equibalanced	Homer was irritated by Marge.
Target	Disturb	Equibalanced	The boy was disturbed by the girl.
Target	Disturb	Equibalanced	Dora was disturbed by Boots.
Target	Disturb	Equibalanced	The doctor was disturbed by the nurse.
Target	Disturb	Equibalanced	Piglet was disturbed by Tigger.
Target	Anger	Equibalanced	The king was angered by the queen.
Target	Anger	Equibalanced	Lisa was angered by Bart.
Target	Anger	Equibalanced	The doctor was angered by the nurse.
Target	Anger	Equibalanced	Wendy was angered by Bob.
Target	Comfort	Equibalanced	The king was conforted by the queen.
Target	Comfort	Equibalanced	Lisa was comforted by Bart.
Target	Comfort	Equibalanced	The teacher was comforted by the student.
Target	Comfort	Equibalanced	Piglet was comforted by Tigger.
Target	Infuriate	Equibalanced	The prince was infuriated by the princess.
Target	Infuriate	Equibalanced	Dora was infuriated by Bob.
Target	Infuriate	Equibalanced	The boy was infuriated by the girl.
Target	Infuriate	Equibalanced	Homer was infuriated by Marge.
Filler			Boots was flying.
Filler			The princess jumped.
Filler			Piglet and Tigger bounced.
Filler			The king and queen waved.
Filler			Tigger was washing.
Filler			The prince was rocking.
Filler			Piglet waved.
Filler			The cat was swinging.
Filler			Dora was flying.
Filler			Bob was swinging.
Filler			The princess and the cat were rocking.

Filler	Dora and Boots waved.
Filler	Bob was flying.
Filler	The prince jumped.
Filler	Tigger was rocking.
Filler	The queen waved.
Filler	The king and queen bounced.
Filler	Piglet jumped.
Filler	Wendy was flying.
Filler	Dora was washing.
Filler	The boy waved.
Filler	Boots pointed at Dora.
Filler	Wendy and Bob jumped.
Filler	Dora was swinging.
Filler	The girl waved.
Filler	Wendy pointed at Bob.
Filler	Boots was washing.
Filler	Piglet was rocking.
Filler	The cat bounced.
Filler	Bob jumped.
Filler	Boots waved at Dora and the baby.
Filler	The king pointed at the queen.
Filler	The doctor and the nurse waved.
Filler	The mouse was rocking.
Filler	The student pointed at the teacher.
Filler	The student was swinging.
Filler	Homer was washing.
Filler	Homer and Marge bounced.
Filler	Lisa and the fox were flying.
Filler	Bart jumped.

6.2. Sentence lists for Study 2

6.2.1. Sentence lists for Study 2 – pilot study

	Verb	**Bias**	**Sentences (DOD version)**
Prime	Bring	PD biased	The prince brought the princess a baby.
Prime	Bring	PD biased	Wendy brought Bob a bird.
Prime	Send	PD biased	The girl sent the boy a fish.
Prime	Send	PD biased	Dora sent Boots a rabbit.
Prime	Take	PD biased	The king took the queen a cat.
Prime	Take	PD biased	Piglet took Tigger a puppy.
Prime	Sell	PD biased	The student sold the teacher a monkey.
Prime	Sell	PD biased	Marge sold Homer a butterfly.
Prime	Leave	PD biased	The doctor left the nurse a mouse.
Prime	Leave	PD biased	Lisa left Bart a fox.
Prime	Hand	DOD biased	The king handed the queen a cat.

Prime	Hand	DOD biased	Piglet handed Tigger a puppy.
Prime	Offer	DOD biased	The doctor offered the nurse a mouse.
Prime	Offer	DOD biased	Lisa offered Bart a fox.
Prime	Award	DOD biased	The student awarded the teacher a monkey
Prime	Award	DOD biased	Marge awarded Homer a butterfly.
Prime	Give	DOD biased	The girl gave the boy a fish.
Prime	Give	DOD biased	Dora gave Boots the rabbit.
Prime	Show	DOD biased	The prince showed the princess a baby.
Prime	Show	DOD biased	Wendy showed Bob a bird.
Target	Throw	Equibalanced	The king threw the queen a cat.
Target	Throw	Equibalanced	Wendy threw Bob a fox.
Target	Throw	Equibalanced	The girl threw the boy a fish.
Target	Throw	Equibalanced	Dora threw Boots a rabbit.
Target	Drop	Equibalanced	The prince dropped the princess a baby.
Target	Drop	Equibalanced	Piglet dropped Tigger a puppy.
Target	Drop	Equibalanced	The student dropped the teacher a monkey.
Target	Drop	Equibalanced	Marge dropped Homer a butterfly.
Target	Flick	Equibalanced	The king flicked the queen a cat.
Target	Flick	Equibalanced	Wendy flicked Bob a fox.
Target	Flick	Equibalanced	The doctor flicked the nurse a mouse.
Target	Flick	Equibalanced	Dora flicked boots a rabbit.
Target	Slide	Equibalanced	The prince slid the princess a baby.
Target	Slide	Equibalanced	Piglet slid Tigger a puppy.
Target	Slide	Equibalanced	The doctor slid the nurse a mouse.
Target	Slide	Equibalanced	Lisa slid Bart a fox.
Target	Feed	Equibalanced	The girl fed the boy a fish.
Target	Feed	Equibalanced	Lisa fed Bar a fox.
Target	Feed	Equibalanced	The student fed the teacher a monkey.
Target	Feed	Equibalanced	Marge fed Homer a butterly.
Filler			Boots was flying.
Filler			The princess jumped.
Filler			Piglet and Tigger bounced.
Filler			The king and queen waved.
Filler			Tigger was washing.
Filler			The prince was rocking.
Filler			Piglet waved.
Filler			The cat was swinging.
Filler			Dora was flying.
Filler			Bob was swinging.
Filler			The princess and the cat were rocking.
Filler			Dora and Boots waved.
Filler			Bob was flying.
Filler			The prince jumped.
Filler			Tigger was rocking.
Filler			The queen waved.
Filler			The king and queen bounced.
Filler			Piglet jumped.

Filler	Wendy was flying.
Filler	Dora was washing.
Filler	The boy waved.
Filler	Boots pointed at Dora.
Filler	Wendy and Bob jumped.
Filler	Dora was swinging.
Filler	The girl waved.
Filler	Wendy pointed at Bob.
Filler	Boots was washing.
Filler	Piglet was rocking.
Filler	The cat bounced.
Filler	Bob jumped.
Filler	Boots waved at Dora and the baby.
Filler	The king pointed at the queen.
Filler	The doctor and the nurse waved.
Filler	The mouse was rocking.
Filler	The student pointed at the teacher.
Filler	The student was swinging.
Filler	Homer was washing.
Filler	Homer and Marge bounced.
Filler	Lisa and the fox were flying.
Filler	Bart jumped.

6.2.2. Sentence lists for Study 2 – main study

Group A - DOD bias						
Trial number		**Sentence type**	**Sentence**	**Verb bias**	**Dative structure**	**Verb**
			Baseline-phase			
1	Experimenter	Filler	Bob was swinging on the swing.	NA	NA	swing
2	Participant	Target	Boots gave … a toy to Dora.	DOD	NA	give
3	Experimenter	Filler	Piglet was playing with the puppy.	NA	NA	play
4	Participant	Target	The student brought … a pencil to the teacher.	PD	NA	bring
5	Experimenter	Filler	Lisa was exercising.	NA	NA	exercise
6	Participant	Target	Wendy threw … a ball to Bob.	equi-biased	NA	throw
7	Experimenter	Filler	The student and teacher were chatting.	NA	NA	chat

8	Participant	Target	The girl fed … a sandwich to the boy.	equi-biased	NA	feed
9	Experimenter	Filler	Boots was washing.	NA	NA	wash
10	Participant	Target	Piglet slid … a peach to Tigger.	equi-biased	NA	slide
11	Experimenter	Filler	The prince and princess were jumping.	NA	NA	jump
12	Participant	Target	Homer brought … an orange to Marge.	PD	NA	bring
13	Experimenter	Filler	Marge and Homer were bouncing.	NA	NA	bounce
14	Participant	Target	Bart fed … a cake to Lisa.	equi-biased	NA	feed
15	Experimenter	Filler	The nurse was jogging.	NA	NA	jog
16	Participant	Target	The king threw … an apple to the queen.	equi-biased	NA	throw
17	Experimenter	Filler	Wendy was flying.	NA	NA	fly
18	Participant	Target	The nurse slid … a book to the doctor.	equi-biased	NA	slide
19	Experimenter	Filler	The girl and the boy were jogging.	NA	NA	jog
20	Participant	Target	The prince gave … a napkin to the princess.	DOD	NA	give
			Test- or bias-phase			
21	Experimenter	Filler	Wendy and Bob were smiling.	NA	NA	smile
22	Participant	Filler	The doctor was … singing.	NA	NA	sing
23	Experimenter	Prime	Bart sent Lisa a cake.	PD	DOD	send
24	Participant	Target	The prince took … a napkin to the princess.	PD	NA	take
25	Experimenter	Filler	Homer and Marge were chatting.	NA	NA	chat
26	Participant	Filler	The mouse was … rocking in the rocking chair.	NA	NA	rock
27	Experimenter	Prime	Boots left a toy to Lisa.	PD	PD	leave
28	Participant	Target	The prince showed … a napkin to the princess.	DOD	NA	show

29	Experimenter	Filler	Wendy was singing with the bird.	NA	NA	sing
30	Participant	Filler	The king and queen were … bouncing.	NA	NA	bounce
31	Experimenter	Prime	The nurse sold the doctor a book.	PD	DOD	sell
32	Participant	Target	Bart sent … a cake to Lisa.	PD	NA	send
33	Experimenter	Filler	The girl was nodding.	NA	NA	nod
34	Participant	Filler	Dora was … swinging on the swing.	NA	NA	swing
35	Experimenter	Prime	The king took an apple to the queen.	PD	PD	take
36	Participant	Target	Wendy handed … a ball to Bob.	DOD	NA	hand
37	Experimenter	Filler	The student was rocking in the rocking chair.	NA	NA	rock
38	Participant	Filler	The monkey was … sneezing.	NA	NA	sneeze
39	Experimenter	Prime	Piglet sent Tigger a peach.	PD	DOD	send
40	Participant	Target	The nurse offered … a book to the doctor.	DOD	NA	offer
41	Experimenter	Filler	Bart was walking with the fox.	NA	NA	walk
42	Participant	Filler	The girl and the boy were … dancing.	NA	NA	dance
43	Experimenter	Prime	Homer left an orange to Marge.	PD	PD	leave
44	Participant	Target	The girl took … a sandwich to the boy.	PD	NA	take
45	Experimenter	Filler	Wendy and Bob were jumping around.	NA	NA	jump
46	Participant	Filler	The boy was waving.	NA	NA	wave
47	Experimenter	Prime	The student sold the teacher a pencil.	PD	DOD	sell
48	Participant	Target	Homer showed … an orange to Marge.	DOD	NA	show
49	Experimenter	Filler	The butterfly was flying on a flower.	NA	NA	fly
50	Participant	Filler	Tigger was … washing.	NA	NA	wash
51	Experimenter	Prime	The girl took a sandwich to the boy.	PD	PD	take
52	Participant	Target	Wendy sent … a ball to Bob.	PD	NA	send
53	Experimenter	Filler	The queen was singing.	NA	NA	sing
54	Participant	Filler	Dora and Boots were … waving.	NA	NA	wave

55	Experimenter	Prime	The prince sent the princess a napkin.	PD	DOD	send
56	Participant	Target	Homer sold … an orange to Marge.	PD	NA	sell
57	Experimenter	Filler	The prince was jumping.	NA	NA	jump
58	Participant	Filler	The puppy was … running around.	NA	NA	run
59	Experimenter	Prime	The king left an apple to the queen.	PD	PD	leave
60	Participant	Target	Boots handed … a toy to Dora.	DOD	NA	hand
61	Experimenter	Filler	The king and the queen were waving.	NA	NA	wave
62	Participant	Filler	The teacher and the monkey were … dancing.	NA	NA	dance
63	Experimenter	Prime	Boots sold Dora a toy.	PD	DOD	sell
64	Participant	Target	The student left … a pencil to the teacher.	PD	NA	leave
65	Experimenter	Filler	The bear was dancing.	NA	NA	dance
66	Participant	Filler	Lisa and the fox were … flying.	NA	NA	fly
67	Experimenter	Prime	Piglet took a peach to Tigger.	PD	PD	take
68	Participant	Target	The girl left … a sandwich to the boy.	DOD	NA	leave
69	Experimenter	Filler	Piglet was rocking in the rocking chair.	NA	NA	rock
70	Participant	Filler	The bird was … sneezing.	NA	NA	sneeze
71	Experimenter	Prime	The girl sent the boy a sandwich.	PD	DOD	send
72	Participant	Target	Piglet awarded … a peach to Tigger.	DOD	NA	award
73	Experimenter	Filler	The prince and the princess were nodding.	NA	NA	nod
74	Participant	Filler	Bart was … jumping.	NA	NA	jump
75	Experimenter	Prime	The student left a pencil to the teacher.	PD	PD	leave
76	Participant	Target	Boots sold … a toy to Dora.	PD	NA	sell
77	Experimenter	Filler	The teacher was singing.	NA	NA	sing
78	Participant	Filler	The king was … exercising.	NA	NA	exercise
79	Experimenter	Prime	Bart sold Lisa a cake.	PD	DOD	sell

80	Participant	Target	The girl offered … a sandwich to the boy.	DOD	NA	offer
81	Experimenter	Filler	Homer was washing.	NA	NA	wash
82	Participant	Filler	The doctor and the nurse were … chatting.	NA	NA	chat
83	Experimenter	Prime	Wendy took a ball to Bob.	PD	PD	take
84	Participant	Target	The king left and apple to the queen.	PD	NA	leave
			Post-test phase			
85	Experimenter	Filler	The cat was swinging on the swing.	NA	NA	swing
86	Participant	Target	Bart gave … a cake to Lisa.	DOD	NA	give
87	Experimenter	Filler	The student was smiling.	NA	NA	smile
88	Participant	Target	The girl brought … a sandwich to the boy.	PD	NA	bring
89	Experimenter	Filler	Marge was singing.	NA	NA	sing
90	Participant	Target	Homer threw … an orange to Marge.	equi-biased	NA	throw
91	Experimenter	Filler	The prince was rocking in the rocking chair.	NA	NA	rock
92	Participant	Target	Piglet fed … a peach to Tigger.	equi-biased	NA	feed
93	Experimenter	Filler	The cat was bouncing.	NA	NA	bounce
94	Participant	Target	The student slid … a pencil to the teacher.	equi-biased	NA	slide
95	Experimenter	Filler	Piglet and Tigger were bouncing.	NA	NA	bounce
96	Participant	Target	Wendy brought … a ball to Bob.	PD	NA	bring
97	Experimenter	Filler	The doctor and the nurse were waving.	NA	NA	wave
98	Participant	Target	The king fed … an apple to the queen.	equi-biased	NA	feed
99	Experimenter	Filler	Homer was playing with the butterfly.	NA	NA	play
100	Participant	Target	The prince threw … a napkin to the princess.	equi-biased	NA	throw
101	Experimenter	Filler	The fox was playing	NA	NA	play

102	Participant	Target	Boots slid … a toy to Dora.	equi-biased	NA	slide
103	Experimenter	Filler	Bart was dancing with Lisa.	NA	NA	dance
104	Participant	Target	The doctor gave … a book to the nurse.	DOD	NA	give
			Second post-test			
105	Experimenter	Filler	Boots was playing with the bunny.	NA	NA	play
106	Participant	Target	Homer sent … Marge an orange.	PD	NA	send
107	Experimenter	Filler	The bunny was dancing.	NA	NA	dance
108	Participant	Target	The prince left … a napkin to the princess.	PD	NA	leave
109	Experimenter	Filler	The prince was playing with a bear.	NA	NA	play
110	Participant	Target	Wendy sold … a Ball to Bob.	PD	NA	sell
111	Experimenter	Filler	Piglet and Tigger were waving.	NA	NA	wave
112	Participant	Target	The nurse took … a book to the doctor.	PD	NA	take
113	Experimenter	Filler	The girl was exercising.	NA	NA	exercise
114	Participant	Target	Bart left … a cake to Lisa.	PD	NA	leave
115	Experimenter	Filler	The fish was swimming.	NA	NA	swim
116	Participant	Target	The king sold … an apple to the queen.	PD	NA	sell
117	Experimenter	Filler	The nurse was playing with the mouse.	NA	NA	play
118	Participant	Target	The student sent … a pencil to the teacher.	PD	NA	send
119	Experimenter	Filler	Dora and boots were nodding.	NA	NA	nod
120	Participant	Target	Boots took … a toy to Dora.	PD	NA	take

Group B - PD bias

Trial number		Sentence type	Animation	Verb bias	Dative structure	Verb
			Baseline-phase			
1	Experimenter	Filler	Bob was swinging on the swing.	NA	NA	swing
2	Participant	Target	Boots gave … a toy to Dora.	DOD	NA	give
3	Experimenter	Filler	Piglet was playing with the puppy.	NA	NA	play
4	Participant	Target	The student brought … a pencil to the teacher.	PD	NA	bring
5	Experimenter	Filler	Lisa was exercising.	NA	NA	exercise
6	Participant	Target	Wendy threw … a ball to Bob.	equi-biased	NA	throw
7	Experimenter	Filler	The student and teacher were chatting.	NA	NA	chat
8	Participant	Target	The girl fed … a sandwich to the boy.	equi-biased	NA	feed
9	Experimenter	Filler	Boots was washing.	NA	NA	wash
10	Participant	Target	Piglet slid … a peach to Tigger.	equi-biased	NA	slide
11	Experimenter	Filler	The prince and princess were jumping.	NA	NA	jump
12	Participant	Target	Homer brought … an orange to Marge.	PD	NA	bring
13	Experimenter	Filler	Marge and Homer were bouncing.	NA	NA	bounce
14	Participant	Target	Bart fed … a cake to Lisa.	equi-biased	NA	feed
15	Experimenter	Filler	The nurse was jogging.	NA	NA	jog
16	Participant	Target	The king threw … an apple to the queen.	equi-biased	NA	throw
17	Experimenter	Filler	Wendy was flying.	NA	NA	fly
18	Participant	Target	The nurse slid … a book to the doctor.	equi-biased	NA	slide
19	Experimenter	Filler	The girl and the boy were jogging.	NA	NA	jog
20	Participant	Target	The prince gave … a napkin to the princess.	DOD	NA	give

Test- or bias-phase						
21	Experimenter	Filler	Wendy and Bob were smiling.	NA	NA	smile
22	Participant	Filler	The doctor was … singing.	NA	NA	sing
23	Experimenter	Prime	Bart handed Lisa a cake.	DOD	DOD	hand
24	Participant	Target	The prince took … a napkin to the princess.	PD	NA	take
25	Experimenter	Filler	Homer and Marge were chatting.	NA	NA	chat
26	Participant	Filler	The mouse was … rocking in the rocking chair.	NA	NA	rock
27	Experimenter	Prime	Boots offered a toy to Lisa.	DOD	PD	offer
28	Participant	Target	The prince showed … a napkin to the princess.	DOD	NA	show
29	Experimenter	Filler	Wendy was singing with the bird.	NA	NA	sing
30	Participant	Filler	The king and queen were … bouncing.	NA	NA	bounce
31	Experimenter	Prime	The nurse awarded the doctor a book.	DOD	DOD	award
32	Participant	Target	Bart sent … a cake to Lisa.	PD	NA	send
33	Experimenter	Filler	The girl was nodding.	NA	NA	nod
34	Participant	Filler	Dora was … swinging on the swing.	NA	NA	swing
35	Experimenter	Prime	The king showed an apple to the queen.	DOD	PD	show
36	Participant	Target	Wendy handed … a ball to Bob.	DOD	NA	hand
37	Experimenter	Filler	The student was rocking in the rocking chair.	NA	NA	rock
38	Participant	Filler	The monkey was … sneezing.	NA	NA	sneeze
39	Experimenter	Prime	Piglet handed Tigger a peach.	DOD	DOD	hand
40	Participant	Target	The nurse offered … a book to the doctor.	DOD	NA	offer
41	Experimenter	Filler	Bart was walking with the fox.	NA	NA	walk
42	Participant	Filler	The girl and the boy were … dancing.	NA	NA	dance
43	Experimenter	Prime	Homer offered an orange to Marge.	DOD	PD	offer
44	Participant	Target	The girl took … a sandwich to the boy.	PD	NA	take

45	Experimenter	Filler	Wendy and Bob were jumping around.	NA	NA	jump
46	Participant	Filler	The boy was … waving.	NA	NA	wave
47	Experimenter	Prime	The student awarded the teacher a pencil.	DOD	DOD	award
48	Participant	Target	Homer showed … an orange to Marge.	DOD	NA	show
49	Experimenter	Filler	The butterfly was flying on a flower.	NA	NA	fly
50	Participant	Filler	Tigger was … washing.	NA	NA	wash
51	Experimenter	Prime	The girl showed a sandwich to the boy.	DOD	PD	show
52	Participant	Target	Wendy sent … a ball to Bob.	PD	NA	send
53	Experimenter	Filler	The queen was singing.	NA	NA	sing
54	Participant	Filler	Dora and Boots were … waving.	NA	NA	wave
55	Experimenter	Prime	The prince handed the princess a napkin.	DOD	DOD	hand
56	Participant	Target	Homer sold … an orange to Marge.	PD	NA	sell
57	Experimenter	Filler	The prince was jumping.	NA	NA	jump
58	Participant	Filler	The puppy was … running around.	NA	NA	run
59	Experimenter	Prime	The king offered an apple to the queen.	DOD	PD	offer
60	Participant	Target	Boots handed … a toy to Dora.	DOD	NA	hand
61	Experimenter	Filler	The king and the queen were waving.	NA	NA	wave
62	Participant	Filler	The teacher and the monkey were … dancing.	NA	NA	dance
63	Experimenter	Prime	Boots awarded Dora a toy.	DOD	DOD	award
64	Participant	Target	The student left … a pencil to the teacher.	PD	NA	leave
65	Experimenter	Filler	The bear was dancing.	NA	NA	dance
66	Participant	Filler	Lisa and the fox were … flying.	NA	NA	fly
67	Experimenter	Prime	Piglet showed a peach to Tigger.	DOD	PD	show
68	Participant	Target	The girl awarded … a sandwich to the boy.	DOD	NA	award
69	Experimenter	Filler	Piglet was rocking in the rocking chair.	NA	NA	rock
70	Participant	Filler	The bird was … sneezing.	NA	NA	sneeze

71	Experimenter	Prime	The girl handed the boy a sandwich.	DOD	DOD	hand
72	Participant	Target	Piglet awarded … a peach to Tigger.	DOD	NA	award
73	Experimenter	Filler	The prince and the princess were nodding.	NA	NA	nod
74	Participant	Filler	Bart was … jumping.	NA	NA	jump
75	Experimenter	Prime	The student offered a pencil to the teacher.	DOD	PD	offer
76	Participant	Target	Boots sold … a toy to Dora.	PD	NA	sell
77	Experimenter	Filler	The teacher was singing.	NA	NA	sing
78	Participant	Filler	The king was … exercising.	NA	NA	exercise
79	Experimenter	Prime	Bart awarded Lisa a cake.	DOD	DOD	award
80	Participant	Target	The girl offered… a sandwich to the boy.	DOD	NA	offer
81	Experimenter	Filler	Homer was washing.	NA	NA	wash
82	Participant	Filler	The doctor and the nurse were … chatting.	NA	NA	chat
83	Experimenter	Prime	Wendy showed a ball to Bob.	DOD	PD	show
84	Participant	Target	The king left … and apple to the queen.	PD	NA	leave
			Post-test phase			
85	Experimenter	Filler	The cat was swinging on the swing.	NA	NA	swing
86	Participant	Target	Bart gave … a cake to Lisa.	DOD	NA	give
87	Experimenter	Filler	The student was smiling.	NA	NA	smile
88	Participant	Target	The girl brought … a sandwich to the boy.	PD	NA	bring
89	Experimenter	Filler	Marge was singing.	NA	NA	sing
90	Participant	Target	Homer threw … an orange to Marge.	equi-biased	NA	throw
91	Experimenter	Filler	The prince was rocking in the rocking chair.	NA	NA	rock
92	Participant	Target	Piglet fed … a peach to Tigger.	equi-biased	NA	feed
93	Experimenter	Filler	The cat was bouncing.	NA	NA	bounce

94	Participant	Target	The student slid … a pencil to the teacher.	equi-biased	NA	slide
95	Experimenter	Filler	Piglet and Tigger were bouncing.	NA	NA	bounce
96	Participant	Target	Wendy brought … a ball to Bob.	PD	NA	bring
97	Experimenter	Filler	The doctor and the nurse were waving.	NA	NA	wave
98	Participant	Target	The king fed … an apple to the queen.	equi-biased	NA	feed
99	Experimenter	Filler	Homer was playing with the butterfly.	NA	NA	play
100	Participant	Target	The prince threw … a napkin to the princess.	equi-biased	NA	throw
101	Experimenter	Filler	The fox was playing	NA	NA	play
102	Participant	Target	Boots slid … a toy to Dora.	equi-biased	NA	slide
103	Experimenter	Filler	Bart was dancing with Lisa.	NA	NA	dance
104	Participant	Target	The doctor gave … a book to the nurse.	DOD	NA	give
			Second post-test			
105	Experimenter	Filler	Boots was playing with the bunny.	NA	NA	play
106	Participant	Target	Homer handed … Marge an orange.	DOD	NA	hand
107	Experimenter	Filler	The bunny was dancing.	NA	NA	dance
108	Participant	Target	The prince offered … a napkin to the princess.	DOD	NA	offer
109	Experimenter	Filler	The prince was playing with a bear.	NA	NA	play
110	Participant	Target	Wendy awarded … a Ball to Bob.	DOD	NA	award
111	Experimenter	Filler	Piglet and Tigger were waving.	NA	NA	wave
112	Participant	Target	The nurse showed … a book to the doctor.	DOD	NA	show
113	Experimenter	Filler	The girl was exercising.	NA	NA	exercise
114	Participant	Target	Bart offered … a cake to Lisa.	DOD	NA	offer
115	Experimenter	Filler	The fish was swimming.	NA	NA	swim

116	Participant	Target	The king awarded … an apple to the queen.	DOD	NA	award
117	Experimenter	Filler	The nurse was playing with the mouse.	NA	NA	play
118	Participant	Target	The student handed … a pencil to the teacher.	DOD	NA	hand
119	Experimenter	Filler	Dora and boots were nodding.	NA	NA	nod
120	Participant	Target	Boots showed … a toy to Dora.	DOD	NA	show

6.3. Sentence lists for Study 3

Group A - PD bias

Trial no	Sentence type	Sentence no	Sentence
		Pre-test phase	
1	**F**	**14**	The dancer and the gardener were singing.
2	**T**	**4**	The fairy mailed ... the pigeon to a sloth.
3	**F**	**52**	The explorer was playing.
4	**F**	**67**	The lion and the panther were having dinner.
5	**F**	**235**	The eagle and the penguin were talking on the phone.
6	**T**	**8**	The explorer loaned ... the turkey to a peacock.
7	**F**	**48**	The seal and the ray were chatting.
8	**F**	**6**	The cheetah and the scorpion were riding a bike.
9	**F**	**34**	The lizard was taking a bath.
10	**T**	**6**	The musician threw ... the hyena to a giraffe.
11	**F**	**183**	The seahorse and the pig were skateboarding.
12	**F**	**199**	The cleaner and the carpenter were catching a bus.
13	**F**	**228**	The pigeon and the tiger were camping.
14	**T**	**10**	The nurse handed ... the spider to a crocodile.
15	**F**	**216**	The owl and the dolphin were juggling.
16	**F**	**18**	The guard and the donkey were singing.
17	**F**	**145**	The mole and the hippo were cleaning.
18	**T**	**9**	Mark handed ... the seahorse to a panther.
19	**F**	**28**	The fireman was dancing.
20	**F**	**173**	The moth and the snail were flying a kite.
21	**F**	**62**	George and Emma were having dinner.
22	**T**	**1**	Maria assigned ... the dove to a raccoon.
23	**F**	**185**	The snake and the cow were skateboarding.
24	**F**	**24**	The repairman and the nurse were dancing.
25	**F**	**22**	The flamingo and the ostrich were dancing.
26	**T**	**2**	The pirate assigned ... the crab to a seagull.
27	**F**	**64**	The repairman and the nurse were having dinner.

28	**F**	**237**	The crocodile and the zebra were talking on the phone.
29	**F**	**40**	The duck was taking a bath.
30	**T**	**7**	Michelle loaned ... the mosquito to a frog.
31	**F**	**74**	Linda and Maria were cooking.
32	**F**	**4**	The bird and the dove were riding a bike.
33	**F**	**107**	The panda was shopping.
34	**T**	**3**	Paul mailed ... the camel to a bull.
35	**F**	**37**	The monkey was taking a bath.
36	**F**	**119**	The mosquito and the bee were reading a book.
37	**F**	**33**	The ant and the cow were taking a bath.
38	**T**	**5**	Laura threw ... the lizard to an ostrich.
39	**F**	**180**	The pirate was flying a kite.
40	**F**	**20**	Madison and the queen were singing.
			Bias-/Test phase
41	**F**	**109**	John was doing shopping.
42	**F**	**89**	Chloe was playing tennis.
43	**P**	**17**	The gardener gave the friendly fox to a horse.
44	**T**	**28**	The actress served ... the crocodile a bull
45	**F**	**80**	The painter was cooking.
46	**F**	**46**	The ostrich and the flamingo were chatting.
47	**P**	**27**	Jacob served the friendly penguin to a badger.
48	**T**	**46**	The bandit tossed ... the butterfly to a pigeon.
49	**F**	**217**	The parrot and the spider were juggling.
50	**F**	**137**	The cat and the dove were painting.
51	**P**	**1**	Emily issued the friendly **robin** to a donkey.
52	**T**	**6**	The fairy slid ... the elephant a scorpion.
53	**F**	**31**	The mosquito and the bat were taking a bath.
54	**F**	**45**	The donkey and the guard were chatting.
55	**P**	**6**	The barista slid the friendly bee to a squirrel.
56	**T**	**26**	Maria served ... the eagle an ostrich.
57	**F**	**39**	The hippo and the bull were taking a bath.
58	**F**	**134**	The bee and the mosquito were painting.
59	**P**	**11**	The secretary provided the friendly lion to a deer.
60	**T**	**79**	Anthony recommended ... the lizard to a snail.
61	**F**	**163**	The robin and the swan were picking flowers.
62	**F**	**30**	The barista and the bandit were dancing.
63	**P**	**79**	The prince promised the friendly horse a cow.
64	**T**	**43**	The builder sent ... the turkey to a sloth.
65	**F**	**56**	The pig and the kangaroo were playing.
66	**F**	**181**	The seagull and the rhino were skateboarding.
67	**P**	**64**	The repairman slipped the friendly moth a bat.
68	**T**	**13**	The explorer provided ... the frog a fish.
69	**F**	**166**	The badger and the rooster were picking flowers.
70	**F**	**136**	Ethan and the plumber were painting.
71	**P**	**26**	The waiter served the friendly monkey to a lion.

72	**T**	**33**	The painter offered ... the owl a giraffe.
73	**F**	**47**	The king and Ella were chatting.
74	**F**	**161**	The rhino and the seagull were picking flowers.
75	**P**	**33**	The knight offered the friendly rhino to a sheep.
76	**T**	**52**	The nurse delivered ... the dove to a spider.
77	**F**	**36**	The seahorse was taking a bath.
78	**F**	**27**	The horse and the camel were dancing.
79	**P**	**18**	Olivia gave the friendly badger to a seal.
80	**T**	**1**	Linda issued ... the bull a seagull.
81	**F**	**139**	The princess was painting.
82	**F**	**164**	The turkey and the squirrel were picking flowers.
83	**P**	**10**	James slid the friendly parrot to a rabbit.
84	**T**	**44**	The carpenter sent ... the mole to a snake.
85	**F**	**82**	The policeman and the princess were playing tennis.
86	**F**	**86**	The turkey and the sheep were playing tennis.
87	**P**	**37**	Sophia promised the friendly squirrel to a dog.
88	**T**	**35**	Paul offered ... the scorpion a panther.
89	**F**	**98**	The rooster was watching tv.
90	**F**	**17**	The magician was singing.
91	**P**	**41**	Aiden issued the friendly **hippo** a rhino.
92	**T**	**64**	Jennifer passed ... the ostrich to a butterfly.
93	**F**	**178**	The assistant and the actress were flying a kite.
94	**F**	**84**	The camel and the horse were playing tennis.
95	**P**	**35**	Michael offered the friendly horse to a tiger.
96	**T**	**23**	John slipped ... the lizard a rooster.
97	**F**	**42**	The dove and the bird were chatting.
98	**F**	**239**	The seagull and the robin were talking on the phone.
99	**P**	**46**	The policeman slid the friendly bear a cheetah.
100	**T**	**76**	The fairy recommended ... the sloth to a mosquito.
101	**F**	**123**	The hedgehog and the sloth were hiking.
102	**F**	**176**	Michael and Matthew were flying a kite.
103	**P**	**77**	The witch promised the friendly rabbit a badger.
104	**T**	**7**	George slid ... the fish a ray.
105	**F**	**7**	The chick and the snake were riding a bike.
106	**F**	**26**	Ella and the king were dancing.
107	**P**	**29**	Chloe served the friendly rat to a robin.
108	**T**	**25**	Sandra slipped ... the fly a swan.
109	**F**	**2**	The bat and the mosquito were riding a bike.
110	**F**	**113**	The rabbit was reading a book.
111	**P**	**59**	The waiter gave the friendly flamingo a penguin.
112	**T**	**8**	Jennifer slid ... the panther a hyena.
113	**F**	**127**	The chick and the scorpion were hiking.
114	**F**	**152**	Olivia and Noah were playing the guitar.
115	**P**	**9**	Isabella slid the friendly donkey to a bear.
116	**T**	**16**	Laura gave ... the turkey a octopus.
117	**F**	**58**	Abigail and Aiden were playing.

118	**F**	**141**	Jennifer and Helen were cleaning.
119	**P**	**71**	Ella offered the friendly pig a goat.
120	**T**	**60**	Maria donated ... the ray to a giraffe.
121	**F**	**77**	The mime was cooking.
122	**F**	**191**	The pigeon and the squirrel were catching a bus.
123	**P**	**68**	The assistant served the friendly wolf a hippo
124	**T**	**15**	Mark provided ... the snake a camel.
125	**F**	**10**	The penguin was riding a bike.
126	**F**	**211**	The swan and the robin were juggling.
127	**P**	**63**	Jacob slipped the friendly deer a zebra.
128	**T**	**77**	The superhero recommended ... the pigeon to an owl.
129	**F**	**116**	The squirrel and the turkey were reading a book.
130	**F**	**90**	The goat and the shark were playing tennis.
131	**P**	**20**	Abigail gave the friendly seal to a monkey.
132	**T**	**41**	Sandra sent ... the swan to a panda.
133	**F**	**121**	The lion and the panther were hiking.
134	**F**	**125**	The builder was hiking.
135	**P**	**4**	The doctor issued the friendly hedgehog to a chick.
136	**T**	**73**	The dancer presented ... the seagull to a turkey.
137	**F**	**188**	William was skateboarding.
138	**F**	**208**	The ray was jumping up and down.
139	**P**	**49**	The student slid the friendly monkey a koala.
140	**T**	**31**	Helen offered ... the kangaroo a sloth.
141	**F**	**169**	The snail and the mouse were picking flowers.
142	**F**	**131**	The peacock and the toucan were painting.
143	**P**	**62**	The knight slipped the friendly seal a turtle.
144	**T**	**19**	The solider gave ... the ostrich a seahorse.
145	**F**	**196**	The dove and the cat were catching a bus.
146	**F**	**110**	The deer and the antelope were doing shopping.
147	**P**	**13**	Madison provided the friendly pelican to a fox.
148	**T**	**29**	The builder served ... the seagull a duck.
149	**F**	**158**	The lion and the fly were playing the guitar.
150	**F**	**223**	The lizard and the dog were camping.
151	**P**	**43**	The fireman issued the friendly bird a deer.
152	**T**	**55**	Susan delivered ... the bull to a scorpion.
153	**F**	**5**	The cat and the octopus were riding a bike.
154	**F**	**143**	The butterfly and the bear were cleaning.
155	**P**	**69**	The cook served the friendly rhino a rabbit.
156	**T**	**5**	Charles issued ... the ant an owl.
157	**F**	**12**	The bee and the butterfly were singing.
158	**F**	**71**	The cowboy and the cowgirl were cooking.
159	**P**	**50**	Chloe slid the friendly badger a pelican.
160	**T**	**68**	The musician sold ... the snail to a dove.
161	**F**	**202**	The ant and the deer were jumping up and down.
162	**F**	**103**	Patricia was doing shopping.
163	**P**	**44**	Olivia issued the friendly penguin a rat.

164	**T**	**37**	The superhero promised ... the duck a raccoon.
165	**F**	**231**	The turtle and the frog were talking on the phone.
166	**F**	**93**	The hedgehog and the mole were watching tv.
167	**P**	**75**	Sophia offered the friendly cheetah a lion.
168	**T**	**78**	Barbara recommended ... the elephant to a frog.
169	**F**	**175**	The pig and the seahorse were flying a kite.
170	**F**	**193**	The hamster and rhino were catching a bus.
171	**P**	**54**	Emma provided the friendly hedgehog a wolf.
172	**T**	**45**	Laura sent ... the hyena to a kangaroo.
173	**F**	**209**	The tiger and the pigeon were jumping up and down.
174	**F**	**29**	The pelican and the eagle were dancing.
175	**P**	**25**	Ella slipped the friendly dog to a wolf.
176	**T**	**38**	Thomas promised ... the spider a shark.
177	**F**	**233**	The bull and the horse were talking on the phone.
178	**F**	**171**	The rat and the hamster were flying a kite.
179	**P**	**3**	The cook issued the friendly sheep to a goat.
180	**T**	**56**	The guard donated ... the fly to a crab.
181	**F**	**11**	The raccoon and the badger were singing.
182	**F**	**190**	The rat and the duck were skateboarding.
183	**P**	**22**	The princess slipped the friendly chick to a rat.
184	**T**	**48**	The cleaner tossed ... the duck to a octopus.
185	**F**	**186**	The dolphin and the owl were skateboarding.
186	**F**	**19**	The panda and Michelle were singing.
187	**P**	**28**	The queen served the friendly moth to a bird.
188	**T**	**4**	Joseph issued ... the seahorse a hamster.
189	**F**	**189**	The spider and the parrot were skateboarding.
190	**F**	**154**	The musician was playing the guitar.
191	**P**	**57**	The king gave the friendly tiger a bear.
192	**T**	**66**	The solider sold ... the giraffe to a panther.
193	**F**	**197**	The assistant and the actress were catching a bus.
194	**F**	**192**	The snail and the moth were catching a bus.
195	**P**	**61**	Elizabeth slipped the friendly donkey a sheep.
196	**T**	**3**	The cleaner issued ... the dove a lizard.
197	**F**	**234**	The plumber and Ethan were talking on the phone.
198	**F**	**205**	The carpenter and the cleaner were jumping up and down.
199	**P**	**51**	The magician provided the friendly toucan a parrot.
200	**T**	**71**	The actress presented ... the spider to a rooster.
201	**F**	**220**	The peacock was juggling.
202	**F**	**194**	The fly and the lion were catching a bus.
203	**P**	**72**	The teacher offered the friendly bat a bee.
204	**T**	**14**	Michelle provided ... the mole a crocodile.
205	**F**	**201**	The spider and the mouse were jumping up and down.
206	**F**	**149**	The fish was cleaning.
207	**P**	**2**	The king issued the friendly cheetah to a pelican.
208	**T**	**20**	The mime gave ... the sloth a pigeon.

209	**F**	**203**	The crab and the pelican were jumping up and down.
210	**F**	**214**	The ostrich and the fox were juggling.
211	**P**	**55**	Matthew provided the friendly chick a horse.
212	**T**	**40**	The bandit promised ... the hamster a peacock.
213	**F**	**21**	The emperor was dancing.
214	**F**	**1**	The badger and the raccoon were riding a bike.
215	**P**	**78**	William promised the friendly mouse a fox.
216	**T**	**49**	Helen tossed ... the panda to an elephant.
217	**F**	**240**	The sheep was talking on the phone.
218	**F**	**51**	The scorpion and the cheetah were playing.
219	**P**	**66**	Abigail served the friendly pelican a flamingo.
220	**T**	**80**	Michelle recommended ... the frog to a seahorse.
221	**F**	**215**	The raccoon and the bat were juggling.
222	**F**	**151**	The cook was playing the guitar.
223	**P**	**58**	Lisa gave the friendly zebra a donkey.
224	**T**	**75**	Charles presented ... the camel to a hyena.
225	**F**	**218**	The kangaroo and the rabbit were juggling.
226	**F**	**226**	The robin and the seagull were camping.
227	**P**	**53**	The plumber provided the friendly turtle a dolphin.
228	**T**	**9**	The pirate slid ... the giraffe a fly.
229	**F**	**115**	Laura was reading a book.
230	**F**	**222**	The scorpion and the chick were camping.
231	**P**	**67**	Noah served the friendly cat a mouse.
232	**T**	**59**	Linda donated ... the fish to a mole.
233	**F**	**16**	Emma and George were singing.
234	**F**	**238**	The turtle was talking on the phone.
235	**P**	**80**	The barista promised the friendly rat a toucan.
236	**T**	**12**	The cowboy provided ... the swan a snake.
237	**F**	**43**	The gardener and the dancer were chatting.
238	**F**	**32**	The butterfly and the bee were taking a bath.
239	**P**	**23**	The teacher slipped the friendly cow to a hippo.
240	**T**	**53**	George delivered ... the eagle to a shark.
241	**F**	**182**	The elephant and the bird were skateboarding.
242	**F**	**179**	The bat and the raccoon were flying a kite.
243	**P**	**5**	Ethan issued the friendly tiger to a pig.
244	**T**	**50**	Richard tossed ... the crab to a duck.
245	**F**	**106**	The toucan and the peacock were doing shopping.
246	**F**	**69**	The tiger and the hyena were having dinner.
247	**P**	**73**	The secretary offered the friendly dog a monkey.
248	**T**	**39**	The guard promised ... the camel an ant.
249	**F**	**100**	The donkey was watching tv.
250	**F**	**15**	The snake and the chick were singing.
251	**P**	**60**	The princess gave the friendly goat a squirrel.
252	**T**	**32**	The singer offered ... the mosquito a butterfly.
253	**F**	**96**	The giraffe was watching tv.
254	**F**	**23**	The fox and the shark were dancing.

255	**P**	**76**	Isabella promised the friendly parrot a bird.
256	**T**	**61**	The explorer passed ... the kangaroo to a fish.
257	**F**	**60**	The koala and the fly were playing.
258	**F**	**138**	Matthew and Michael were painting.
259	**P**	**52**	The gardener provided the friendly robin a chick.
260	**T**	**36**	Helen promised ... the crab a dove.
261	**F**	**168**	The fox and the ostrich were picking flowers.
262	**F**	**41**	The elephant and the bear were chatting.
263	**P**	**70**	The doctor served the friendly fox a pig.
264	**T**	**74**	Paul presented ... the crocodile to a peacock.
265	**F**	**54**	The shark and the fox were playing.
266	**F**	**55**	The sloth and the goat were playing.
267	**P**	**24**	Daniel slipped the friendly flamingo to a bee.
268	**T**	**70**	Mark sold ... the rooster to an ant.
269	**F**	**65**	The mole and the hedgehog were having dinner.
270	**F**	**95**	The ray and the seal were watching tv.
271	**P**	**12**	The repairman provided the friendly bird to a mouse.
272	**T**	**69**	Patricia sold ... the hamster to a swan.
273	**F**	**81**	The cowgirl and the cowboy were playing tennis.
274	**F**	**135**	The octopus was painting.
275	**P**	**45**	Michael issued the friendly cow a dog.
276	**T**	**24**	The nurse slipped ... the snail a mosquito.
277	**F**	**156**	The cheetah was playing the guitar.
278	**F**	**162**	The knight and the magician were picking flowers.
279	**P**	**48**	Daniel slid the friendly lion a tiger.
280	**T**	**58**	Joseph donated ... the raccoon to a ray.
281	**F**	**148**	The witch was cleaning.
282	**F**	**167**	The cow and the snake were picking flowers.
283	**P**	**16**	The prince gave the friendly koala to a parrot.
284	**T**	**21**	Susan slipped ... the panda a turkey.
285	**F**	**88**	Noah and Olivia were playing tennis.
286	**F**	**78**	Sophia and the waiter were cooking.
287	**P**	**36**	The policeman promised the friendly goat to a bat.
288	**T**	**30**	The carpenter served ... the shark a kangaroo.
289	**F**	**59**	Helen and Jennifer were playing.
290	**F**	**105**	The doctor was doing shopping.
291	**P**	**21**	The plumber slipped the friendly pig to a dolphin.
292	**T**	**10**	Patricia slid ... the raccoon an eagle.
293	**F**	**79**	Aiden and Abigail were cooking.
294	**F**	**44**	The cow and the ant were chatting.
295	**P**	**14**	William provided the friendly bat to a flamingo.
296	**T**	**2**	The musician issued ... the butterfly a snail.
297	**F**	**104**	The waiter and Sophia were doing shopping.
298	**F**	**91**	The dog and the owl were watching tv.
299	**P**	**7**	The magician slid the friendly toucan to a rhino.
300	**T**	**63**	The singer passed ... the ant to a bull.

301	**F**	**140**	Susan was painting.
302	**F**	**63**	Richard was having dinner.
303	**P**	**34**	The witch offered the friendly wolf to a toucan.
304	**T**	**18**	Barbara gave ... the rooster a panda.
305	**F**	**66**	The kangaroo and the pig were having dinner.
306	**F**	**38**	The goat and the sloth were taking a bath.
307	**P**	**65**	Emily slipped the friendly koala a robin.
308	**T**	**57**	The cowboy donated ... the shark to a camel.
309	**F**	**146**	Mark was cleaning.
310	**F**	**147**	Emily was cleaning.
311	**P**	**19**	Aiden gave the friendly dolphin to a hedgehog.
312	**T**	**65**	Helen passed ... the owl to a crocodile.
313	**F**	**3**	The bear and the elephant were riding a bike.
314	**F**	**122**	The zebra was hiking.
315	**P**	**40**	Emma promised the friendly turtle to a penguin.
316	**T**	**51**	The painter delivered ... the panther to a lizard.
317	**F**	**25**	Maria and Linda were dancing.
318	**F**	**53**	Michelle and the panda were playing.
319	**P**	**47**	Madison slid the friendly dolphin a seal.
320	**T**	**22**	The dancer slipped ... the ray a frog.
321	**F**	**83**	The queen and Madison were playing tennis.
322	**F**	**35**	The dolphin and the fish were taking a bath.
323	**P**	**39**	Noah promised the friendly deer to a moth.
324	**T**	**72**	The pirate presented ... the octopus to a fly.
325	**F**	**57**	The giraffe was playing.
326	**F**	**49**	Anthony and Barbara were chatting.
327	**P**	**31**	The assistant offered the friendly mouse to a zebra.
328	**T**	**42**	The cowgirl sent ... the scorpion to a eagle.
329	**F**	**68**	The koala and the fly were having dinner.
330	**F**	**73**	The giraffe was cooking.
331	**P**	**8**	The fireman slid the friendly rabbit to a cat.
332	**T**	**11**	The emperor provided ... the peacock a crab.
333	**F**	**70**	Joseph was having dinner.
334	**F**	**165**	The monkey and the panther were picking flowers.
335	**P**	**30**	Matthew served the friendly zebra to a cow.
336	**T**	**67**	John sold ... the seahorse to a raccoon.
337	**F**	**87**	The singer was playing tennis.
338	**F**	**101**	Raul and Paul were doing shopping.
339	**P**	**38**	Lisa promised the friendly bear to a turtle.
340	**T**	**17**	The cowgirl gave ... the hyena an elephant.
341	**F**	**132**	The student and the teacher were painting.
342	**F**	**75**	The bull and the hippo were cooking.
343	**P**	**56**	Ethan gave the friendly sheep a cat.
344	**T**	**54**	Thomas delivered ... the mosquito to an ostrich.
345	**F**	**129**	The dog and the lizard were hiking.
346	**F**	**76**	Barbara and Anthony were cooking.

347	**P**	**74**	James offered the friendly squirrel a hedgehog.
348	**T**	**27**	Anthony served ... the octopus a spider.
349	**F**	**170**	Elizabeth was picking flowers.
350	**F**	**157**	The bird and the elephant were playing the guitar.
351	**P**	**42**	The queen issued a friendly bee a moth.
352	**T**	**34**	Richard offered ... the pigeon a mole.
353	**F**	**229**	The frog and the wolf were camping.
354	**F**	**108**	The shark and the goat were doing shopping.
355	**P**	**32**	The student offered the friendly cat to a koala.
356	**T**	**62**	The mime passed ... the snake to a hamster.
357	**F**	**92**	The fish and the dolphin were watching tv.
358	**F**	**117**	The camel and the koala were reading a book.
359	**P**	**15**	Elizabeth provided the friendly hippo to a cheetah.
360	**T**	**47**	The emperor tossed ... the peacock to a seagull.
		Post-test phase	
361	**F**	**50**	The mouse and the spider were chatting.
362	**F**	**99**	The crocodile and the wolf were watching tv.
363	**T**	**20**	The solider awarded the... the scorpion to a kangaroo.
364	**F**	**111**	The wolf and the crocodile were reading a book.
365	**F**	**114**	The teacher and the student were reading a book.
366	**F**	**172**	The squirrel and the pigeon were flying a kite.
367	**T**	**17**	Sandra brought the... the fly to an octopus.
368	**F**	**224**	The penguin and the eagle were camping.
369	**F**	**130**	Charles was hiking.
370	**F**	**126**	The hippo and the mole were hiking.
371	**T**	**18**	The guard brought the... the butterfly to a snail.
372	**F**	**112**	The sheep and the turkey were reading a book.
373	**F**	**219**	The zebra and the crocodile were juggling.
374	**F**	**94**	The eagle and the pelican were watching tv.
375	**T**	**13**	Anthony lent the... the ant to a panda.
376	**F**	**97**	The hamster and the rat were watching tv.
377	**F**	**128**	Lisa and Daniel were hiking.
378	**F**	**102**	The owl and the dog were doing shopping.
379	**T**	**14**	The carpenter lent the... the mole to a shark.
380	**F**	**225**	The giraffe and the moth were camping.
381	**F**	**94**	The eagle and the pelican were watching tv.
382	**F**	**210**	The wolf and the frog were jumping up and down.
383	**T**	**15**	George showed the... the duck to an eagle.
384	**F**	**72**	Paul and Raul were cooking.
385	**F**	**142**	The horse and the bull were cleaning.
386	**F**	**184**	The panther and the monkey were skateboarding.
387	**T**	**19**	Helen awarded the... the snake to a rooster.
388	**F**	**153**	The fairy was playing the guitar.
389	**F**	**120**	Isabella was reading a book.
390	**F**	**144**	The seal and the swan were cleaning.

391	**T**	**16**	The cowgirl showed the... the fish to an elephant.
392	**F**	**227**	The hyena and the toucan were camping.
393	**F**	**124**	The pelican and the crab were hiking.
394	**F**	**61**	The ant and cow were having dinner.
395	**T**	**12**	The actress faxed the... the hamster to a swan.
396	**F**	**118**	The bear and the butterfly were reading a book.
397	**F**	**160**	Sandra was playing the guitar.
398	**F**	**133**	The rhino and the hamster were painting.
399	**T**	**11**	Charles faxed the... the owl to a ray.
400	**F**	**85**	The bandit and the barista were playing tennis.

Group B - DOD bias

Trial no	Sentence type	Sentence no	Sentence
			Pre-test phase
1	**F**	**14**	The dancer and the gardener were singing.
2	**T**	**4**	The fairy mailed ... the pigeon to a sloth.
3	**F**	**52**	The explorer was playing.
4	**F**	**67**	The lion and the panther were having dinner.
5	**F**	**235**	The eagle and the penguin were talking on the phone.
6	**T**	**8**	The explorer loaned ... the turkey to a peacock.
7	**F**	**48**	The seal and the ray were chatting.
8	**F**	**6**	The cheetah and the scorpion were riding a bike.
9	**F**	**34**	The lizard was taking a bath.
10	**T**	**6**	The musician threw ... the hyena to a giraffe.
11	**F**	**183**	The seahorse and the pig were skateboarding.
12	**F**	**199**	The cleaner and the carpenter were catching a bus.
13	**F**	**228**	The pigeon and the tiger were camping.
14	**T**	**10**	The nurse handed ... the spider to a crocodile.
15	**F**	**216**	The owl and the dolphin were juggling.
16	**F**	**18**	The guard and the donkey were singing.
17	**F**	**145**	The mole and the hippo were cleaning.
18	**T**	**9**	Mark handed ... the seahorse to a panther.
19	**F**	**28**	The fireman was dancing.
20	**F**	**173**	The moth and the snail were flying a kite.
21	**F**	**62**	George and Emma were having dinner.
22	**T**	**1**	Maria assigned ... the dove to a raccoon.
23	**F**	**185**	The snake and the cow were skateboarding.
24	**F**	**24**	The repairman and the nurse were dancing.
25	**F**	**22**	The flamingo and the ostrich were dancing.
26	**T**	**2**	The pirate assigned ... the crab to a seagull.
27	**F**	**64**	The repairman and the nurse were having dinner.
28	**F**	**237**	The crocodile and the zebra were talking on the phone.
29	**F**	**40**	The duck was taking a bath.

30	**T**	**7**	Michelle loaned ... the mosquito to a frog.
31	**F**	**74**	Linda and Maria were cooking.
32	**F**	**4**	The bird and the dove were riding a bike.
33	**F**	**107**	The panda was shopping.
34	**T**	**3**	Paul mailed ... the camel to a bull.
35	**F**	**37**	The monkey was taking a bath.
36	**F**	**119**	The mosquito and the bee were reading a book.
37	**F**	**33**	The ant and the cow were taking a bath.
38	**T**	**5**	Laura threw ... the lizard to an ostrich.
39	**F**	**180**	The pirate was flying a kite.
40	**F**	**20**	Madison and the queen were singing.
			Bias-/Test phase
41	**F**	**109**	John was doing shopping.
42	**F**	**89**	Chloe was playing tennis.
43	**P**	**17**	The gardener donated the friendly **fox** to a horse.
44	**T**	**28**	The actress served ... the crocodile a bull
45	**F**	**80**	The painter was cooking.
46	**F**	**46**	The ostrich and the flamingo were chatting.
47	**P**	**27**	Jacob sold the friendly **penguin** to a badger.
48	**T**	**46**	The bandit tossed ... the butterfly to a pigeon.
49	**F**	**217**	The parrot and the spider were juggling.
50	**F**	**137**	The cat and the dove were painting.
51	**P**	**1**	Emily sent the friendly **robin** to a donkey.
52	**T**	**6**	The fairy slid ... the elephant a scorpion.
53	**F**	**31**	The mosquito and the bat were taking a bath.
54	**F**	**45**	The donkey and the guard were chatting.
55	**P**	**6**	The barista tossed the friendly **bee** to a squirrel.
56	**T**	**26**	Maria served ... the eagle an ostrich.
57	**F**	**39**	The hippo and the bull were taking a bath.
58	**F**	**134**	The bee and the mosquito were painting.
59	**P**	**11**	The secretary delivered the friendly **lion** to a deer.
60	**T**	**79**	Anthony recommended ... the lizard to a snail.
61	**F**	**163**	The robin and the swan were picking flowers.
62	**F**	**30**	The barista and the bandit were dancing.
63	**P**	**79**	The prince recommended the friendly **horse** a cow.
64	**T**	**43**	The builder sent ... the turkey to a sloth.
65	**F**	**56**	The pig and the kangaroo were playing.
66	**F**	**181**	The seagull and the rhino were skateboarding.
67	**P**	**64**	The repairman passed the friendly **moth** a bat.
68	**T**	**13**	The explorer provided ... the frog a fish.
69	**F**	**166**	The badger and the rooster were picking flowers.
70	**F**	**136**	Ethan and the plumber were painting.
71	**P**	**26**	The waiter sold the friendly **monkey** to a lion.
72	**T**	**33**	The painter offered ... the owl a giraffe.
73	**F**	**47**	The king and Ella were chatting.
74	**F**	**161**	The rhino and the seagull were picking flowers.

75	**P**	**33**	The knight presented the friendly **rhino** to a sheep.
76	**T**	**52**	The nurse delivered ... the dove to a spider.
77	**F**	**36**	The seahorse was taking a bath.
78	**F**	**27**	The horse and the camel were dancing.
79	**P**	**18**	Olivia donated the friendly **badger** to a seal.
80	**T**	**1**	Linda issued ... the bull a seagull.
81	**F**	**139**	The princess was painting.
82	**F**	**164**	The turkey and the squirrel were picking flowers.
83	**P**	**10**	James tossed the friendly **parrot** to a rabbit.
84	**T**	**44**	The carpenter sent ... the mole to a snake.
85	**F**	**82**	The policeman and the princess were playing tennis.
86	**F**	**86**	The turkey and the sheep were playing tennis.
87	**P**	**37**	Sophia recommended the friendly **squirrel** to a dog.
88	**T**	**35**	Paul offered ... the scorpion a panther.
89	**F**	**98**	The rooster was watching tv.
90	**F**	**17**	The magician was singing.
91	**P**	**41**	Aiden sent the friendly **hippo** a rhino.
92	**T**	**64**	Jennifer passed ... the ostrich to a butterfly.
93	**F**	**178**	The assistant and the actress were flying a kite.
94	**F**	**84**	The camel and the horse were playing tennis.
95	**P**	**35**	Michael presented the friendly **horse** to a tiger.
96	**T**	**23**	John slipped ... the lizard a rooster.
97	**F**	**42**	The dove and the bird were chatting.
98	**F**	**239**	The seagull and the robin were talking on the phone.
99	**P**	**46**	The policeman tossed the friendly **bear** a cheetah.
100	**T**	**76**	The fairy recommended ... the sloth to a mosquito.
101	**F**	**123**	The hedgehog and the sloth were hiking.
102	**F**	**176**	Michael and Matthew were flying a kite.
103	**P**	**77**	The witch recommended the friendly **rabbit** a badger.
104	**T**	**7**	George slid ... the fish a ray.
105	**F**	**7**	The chick and the snake were riding a bike.
106	**F**	**26**	Ella and the king were dancing.
107	**P**	**29**	Chloe sold the friendly **rat** to a robin.
108	**T**	**25**	Sandra slipped ... the fly a swan.
109	**F**	**2**	The bat and the mosquito were riding a bike.
110	**F**	**113**	The rabbit was reading a book.
111	**P**	**59**	The waiter donated the friendly **flamingo** a penguin.
112	**T**	**8**	Jennifer slid ... the panther a hyena.
113	**F**	**127**	The chick and the scorpion were hiking.
114	**F**	**152**	Olivia and Noah were playing the guitar.
115	**P**	**9**	Isabella tossed the friendly **donkey** to a bear.
116	**T**	**16**	Laura gave ... the turkey a octopus.
117	**F**	**58**	Abigail and Aiden were playing.
118	**F**	**141**	Jennifer and Helen were cleaning.
119	**P**	**71**	Ella presented the friendly **pig** a goat.

120	**T**	**60**	Maria donated ... the ray to a giraffe.
121	**F**	**77**	The mime was cooking.
122	**F**	**191**	The pigeon and the squirrel were catching a bus.
123	**P**	**68**	The assistant sold the friendly **wolf** a hippo
124	**T**	**15**	Mark provided ... the snake a camel.
125	**F**	**10**	The penguin was riding a bike.
126	**F**	**211**	The swan and the robin were juggling.
127	**P**	**63**	Jacob passed the friendly **deer** a zebra.
128	**T**	**77**	The superhero recommended ... the pigeon to an owl.
129	**F**	**116**	The squirrel and the turkey were reading a book.
130	**F**	**90**	The goat and the shark were playing tennis.
131	**P**	**20**	Abigail donated the friendly **seal** to a monkey.
132	**T**	**41**	Sandra sent ... the swan to a panda.
133	**F**	**121**	The lion and the panther were hiking.
134	**F**	**125**	The builder was hiking.
135	**P**	**4**	The doctor sent the friendly **hedgehog** to a chick.
136	**T**	**73**	The dancer presented ... the seagull to a turkey.
137	**F**	**188**	William was skateboarding.
138	**F**	**208**	The ray was jumping up and down.
139	**P**	**49**	The student tossed the friendly **monkey** a koala.
140	**T**	**31**	Helen offered ... the kangaroo a sloth.
141	**F**	**169**	The snail and the mouse were picking flowers.
142	**F**	**131**	The peacock and the toucan were painting.
143	**P**	**62**	The knight passed the friendly **seal** a turtle.
144	**T**	**19**	The solider gave ... the ostrich a seahorse.
145	**F**	**196**	The dove and the cat were catching a bus.
146	**F**	**110**	The deer and the antelope were doing shopping.
147	**P**	**13**	Madison delivered the friendly **pelican** to a fox.
148	**T**	**29**	The builder served ... the seagull a duck.
149	**F**	**158**	The lion and the fly were playing the guitar.
150	**F**	**223**	The lizard and the dog were camping.
151	**P**	**43**	The fireman sent the friendly **bird** a deer.
152	**T**	**55**	Susan delivered ... the bull to a scorpion.
153	**F**	**5**	The cat and the octopus were riding a bike.
154	**F**	**143**	The butterfly and the bear were cleaning.
155	**P**	**69**	The cook sold the friendly **rhino** a rabbit.
156	**T**	**5**	Charles issued ... the ant an owl.
157	**F**	**12**	The bee and the butterfly were singing.
158	**F**	**71**	The cowboy and the cowgirl were cooking.
159	**P**	**50**	Chloe tossed the friendly **badger** a pelican.
160	**T**	**68**	The musician sold ... the snail to a dove.
161	**F**	**202**	The ant and the deer were jumping up and down.
162	**F**	**103**	Patricia was doing shopping.
163	**P**	**44**	Olivia sent the friendly **penguin** a rat.
164	**T**	**37**	The superhero promised ... the duck a raccoon.

165	**F**	**231**	The turtle and the frog were talking on the phone.
166	**F**	**93**	The hedgehog and the mole were watching tv.
167	**P**	**75**	Sophia presented the friendly **cheetah** a lion.
168	**T**	**78**	Barbara recommended ... the elephant to a frog.
169	**F**	**175**	The pig and the seahorse were flying a kite.
170	**F**	**193**	The hamster and rhino were catching a bus.
171	**P**	**54**	Emma delivered the friendly **hedgehog** a wolf.
172	**T**	**45**	Laura sent ... the hyena to a kangaroo.
173	**F**	**209**	The tiger and the pigeon were jumping up and down.
174	**F**	**29**	The pelican and the eagle were dancing.
175	**P**	**25**	Ella passed the friendly **dog** to a wolf.
176	**T**	**38**	Thomas promised ... the spider a shark.
177	**F**	**233**	The bull and the horse were talking on the phone.
178	**F**	**171**	The rat and the hamster were flying a kite.
179	**P**	**3**	The cook sent the friendly **sheep** to a goat.
180	**T**	**56**	The guard donated ... the fly to a crab.
181	**F**	**11**	The raccoon and the badger were singing.
182	**F**	**190**	The rat and the duck were skateboarding.
183	**P**	**22**	The princess passed the friendly **chick** to a rat.
184	**T**	**48**	The cleaner tossed ... the duck to a octopus.
185	**F**	**186**	The dolphin and the owl were skateboarding.
186	**F**	**19**	The panda and Michelle were singing.
187	**P**	**28**	The queen sold the friendly **moth** to a bird.
188	**T**	**4**	Joseph issued ... the seahorse a hamster.
189	**F**	**189**	The spider and the parrot were skateboarding.
190	**F**	**154**	The musician was playing the guitar.
191	**P**	**57**	The king donated the friendly **tiger** a bear.
192	**T**	**66**	The solider sold ... the giraffe to a panther.
193	**F**	**197**	The assistant and the actress were catching a bus.
194	**F**	**192**	The snail and the moth were catching a bus.
195	**P**	**61**	Elizabeth passed the friendly **donkey** a sheep.
196	**T**	**3**	The cleaner issued ... the dove a lizard.
197	**F**	**234**	The plumber and Ethan were talking on the phone.
198	**F**	**205**	The carpenter and the cleaner were jumping up and down.
199	**P**	**51**	The magician delivered the friendly **toucan** a parrot.
200	**T**	**71**	The actress presented ... the spider to a rooster.
201	**F**	**220**	The peacock was juggling.
202	**F**	**194**	The fly and the lion were catching a bus.
203	**P**	**72**	The teacher presented the friendly **bat** a bee.
204	**T**	**14**	Michelle provided ... the mole a crocodile.
205	**F**	**201**	The spider and the mouse were jumping up and down.
206	**F**	**149**	The fish was cleaning.
207	**P**	**2**	The king sent the friendly **cheetah** to a pelican.
208	**T**	**20**	The mime gave ... the sloth a pigeon.

209	**F**	**203**	The crab and the pelican were jumping up and down.
210	**F**	**214**	The ostrich and the fox were juggling.
211	**P**	**55**	Matthew delivered the friendly **chick** a horse.
212	**T**	**40**	The bandit promised ... the hamster a peacock.
213	**F**	**21**	The emperor was dancing.
214	**F**	**1**	The badger and the raccoon were riding a bike.
215	**P**	**78**	William recommended the friendly **mouse** a fox.
216	**T**	**49**	Helen tossed ... the panda to an elephant.
217	**F**	**240**	The sheep was talking on the phone.
218	**F**	**51**	The scorpion and the cheetah were playing.
219	**P**	**66**	Abigail sold the friendly **pelican** a flamingo.
220	**T**	**80**	Michelle recommended ... the frog to a seahorse.
221	**F**	**215**	The raccoon and the bat were juggling.
222	**F**	**151**	The cook was playing the guitar.
223	**P**	**58**	Lisa donated the friendly **zebra** a donkey.
224	**T**	**75**	Charles presented ... the camel to a hyena.
225	**F**	**218**	The kangaroo and the rabbit were juggling.
226	**F**	**226**	The robin and the seagull were camping.
227	**P**	**53**	The plumber delivered the friendly **turtle** a dolphin.
228	**T**	**9**	The pirate slid ... the giraffe a fly.
229	**F**	**115**	Laura was reading a book.
230	**F**	**222**	The scorpion and the chick were camping.
231	**P**	**67**	Noah sold the friendly **cat** a mouse.
232	**T**	**59**	Linda donated ... the fish to a mole.
233	**F**	**16**	Emma and George were singing.
234	**F**	**238**	The turtle was talking on the phone.
235	**P**	**80**	The barista recommended the friendly **rat** a toucan.
236	**T**	**12**	The cowboy provided ... the swan a snake.
237	**F**	**43**	The gardener and the dancer were chatting.
238	**F**	**32**	The butterfly and the bee were taking a bath.
239	**P**	**23**	The teacher passed the friendly **cow** to a hippo.
240	**T**	**53**	George delivered ... the eagle to a shark.
241	**F**	**182**	The elephant and the bird were skateboarding.
242	**F**	**179**	The bat and the raccoon were flying a kite.
243	**P**	**5**	Ethan sent the friendly **tiger** to a pig.
244	**T**	**50**	Richard tossed ... the crab to a duck.
245	**F**	**106**	The toucan and the peacock were doing shopping.
246	**F**	**69**	The tiger and the hyena were having dinner.
247	**P**	**73**	The secretary presented the friendly **dog** a monkey.
248	**T**	**39**	The guard promised ... the camel an ant.
249	**F**	**100**	The donkey was watching tv.
250	**F**	**15**	The snake and the chick were singing.
251	**P**	**60**	The princess donated the friendly **goat** a squirrel.
252	**T**	**32**	The singer offered ... the mosquito a butterfly.
253	**F**	**96**	The giraffe was watching tv.

254	**F**	**23**	The fox and the shark were dancing.
255	**P**	**76**	Isabella recommended the friendly **parrot** a bird.
256	**T**	**61**	The explorer passed ... the kangaroo to a fish.
257	**F**	**60**	The koala and the fly were playing.
258	**F**	**138**	Matthew and Michael were painting.
259	**P**	**52**	The gardener delivered the friendly **robin** a chick.
260	**T**	**36**	Helen promised ... the crab a dove.
261	**F**	**168**	The fox and the ostrich were picking flowers.
262	**F**	**41**	The elephant and the bear were chatting.
263	**P**	**70**	The doctor sold the friendly **fox** a pig.
264	**T**	**74**	Paul presented ... the crocodile to a peacock.
265	**F**	**54**	The shark and the fox were playing.
266	**F**	**55**	The sloth and the goat were playing.
267	**P**	**24**	Daniel passed the friendly **flamingo** to a bee.
268	**T**	**70**	Mark sold ... the rooster to an ant.
269	**F**	**65**	The mole and the hedgehog were having dinner.
270	**F**	**95**	The ray and the seal were watching tv.
271	**P**	**12**	The repairman delivered the friendly **bird** to a mouse.
272	**T**	**69**	Patricia sold ... the hamster to a swan.
273	**F**	**81**	The cowgirl and the cowboy were playing tennis.
274	**F**	**135**	The octopus was painting.
275	**P**	**45**	Michael sent the friendly **cow** a dog.
276	**T**	**24**	The nurse slipped ... the snail a mosquito.
277	**F**	**156**	The cheetah was playing the guitar.
278	**F**	**162**	The knight and the magician were picking flowers.
279	**P**	**48**	Daniel tossed the friendly **lion** a tiger.
280	**T**	**58**	Joseph donated ... the raccoon to a ray.
281	**F**	**148**	The witch was cleaning.
282	**F**	**167**	The cow and the snake were picking flowers.
283	**P**	**16**	The prince donated the friendly **koala** to a parrot.
284	**T**	**21**	Susan slipped ... the panda a turkey.
285	**F**	**88**	Noah and Olivia were playing tennis.
286	**F**	**78**	Sophia and the waiter were cooking.
287	**P**	**36**	The policeman recommended the friendly **goat** to a bat.
288	**T**	**30**	The carpenter served ... the shark a kangaroo.
289	**F**	**59**	Helen and Jennifer were playing.
290	**F**	**105**	The doctor was doing shopping.
291	**P**	**21**	The plumber passed the friendly **pig** to a dolphin.
292	**T**	**10**	Patricia slid ... the raccoon an eagle.
293	**F**	**79**	Aiden and Abigail were cooking.
294	**F**	**44**	The cow and the ant were chatting.
295	**P**	**14**	William delivered the friendly **bat** to a flamingo.
296	**T**	**2**	The musician issued ... the butterfly a snail.
297	**F**	**104**	The waiter and Sophia were doing shopping.
298	**F**	**91**	The dog and the owl were watching tv.

299	**P**	**7**	The magician tossed the friendly **toucan** to a rhino.
300	**T**	**63**	The singer passed ... the ant to a bull.
301	**F**	**140**	Susan was painting.
302	**F**	**63**	Richard was having dinner.
303	**P**	**34**	The witch presented the friendly **wolf** to a toucan.
304	**T**	**18**	Barbara gave ... the rooster a panda.
305	**F**	**66**	The kangaroo and the pig were having dinner.
306	**F**	**38**	The goat and the sloth were taking a bath.
307	**P**	**65**	Emily passed the friendly **koala** a robin.
308	**T**	**57**	The cowboy donated ... the shark to a camel.
309	**F**	**146**	Mark was cleaning.
310	**F**	**147**	Emily was cleaning.
311	**P**	**19**	Aiden donated the friendly **dolphin** to a hedgehog.
312	**T**	**65**	Helen passed ... the owl to a crocodile.
313	**F**	**3**	The bear and the elephant were riding a bike.
314	**F**	**122**	The zebra was hiking.
315	**P**	**40**	Emma recommended the friendly **turtle** to a penguin.
316	**T**	**51**	The painter delivered ... the panther to a lizard.
317	**F**	**25**	Maria and Linda were dancing.
318	**F**	**53**	Michelle and the panda were playing.
319	**P**	**47**	Madison tossed the friendly **dolphin** a seal.
320	**T**	**22**	The dancer slipped ... the ray a frog.
321	**F**	**83**	The queen and Madison were playing tennis.
322	**F**	**35**	The dolphin and the fish were taking a bath.
323	**P**	**39**	Noah recommended the friendly **deer** to a moth.
324	**T**	**72**	The pirate presented ... the octopus to a fly.
325	**F**	**57**	The giraffe was playing.
326	**F**	**49**	Anthony and Barbara were chatting.
327	**P**	**31**	The assistant presented the friendly **mouse** to a zebra.
328	**T**	**42**	The cowgirl sent ... the scorpion to a eagle.
329	**F**	**68**	The koala and the fly were having dinner.
330	**F**	**73**	The giraffe was cooking.
331	**P**	**8**	The fireman tossed the friendly **rabbit** to a cat.
332	**T**	**11**	The emperor provided ... the peacock a crab.
333	**F**	**70**	Joseph was having dinner.
334	**F**	**165**	The monkey and the panther were picking flowers.
335	**P**	**30**	Matthew sold the friendly **zebra** to a cow.
336	**T**	**67**	John sold ... the seahorse to a raccoon.
337	**F**	**87**	The singer was playing tennis.
338	**F**	**101**	Raul and Paul were doing shopping.
339	**P**	**38**	Lisa recommended the friendly **bear** to a turtle.
340	**T**	**17**	The cowgirl gave ... the hyena an elephant.
341	**F**	**132**	The student and the teacher were painting.
342	**F**	**75**	The bull and the hippo were cooking.
343	**P**	**56**	Ethan donated the friendly **sheep** a cat.

344	**T**	**54**	Thomas delivered ... the mosquito to an ostrich.
345	**F**	**129**	The dog and the lizard were hiking.
346	**F**	**76**	Barbara and Anthony were cooking.
347	**P**	**74**	James presented the friendly **squirrel** a hedgehog.
348	**T**	**27**	Anthony served ... the octopus a spider.
349	**F**	**170**	Elizabeth was picking flowers.
350	**F**	**157**	The bird and the elephant were playing the guitar.
351	**P**	**42**	The queen sent a friendly **bee** a moth.
352	**T**	**34**	Richard offered ... the pigeon a mole.
353	**F**	**229**	The frog and the wolf were camping.
354	**F**	**108**	The shark and the goat were doing shopping.
355	**P**	**32**	The student presented the friendly **cat** to a koala.
356	**T**	**62**	The mime passed ... the snake to a hamster.
357	**F**	**92**	The fish and the dolphin were watching tv.
358	**F**	**117**	The camel and the koala were reading a book.
359	**P**	**15**	Elizabeth delivered the friendly **hippo** to a cheetah.
360	**T**	**47**	The emperor tossed ... the peacock to a seagull.
		Post-test phase	
361	**F**	**50**	The mouse and the spider were chatting.
362	**F**	**99**	The crocodile and the wolf were watching tv.
363	**T**	**20**	The solider awarded the... the scorpion to a kangaroo.
364	**F**	**111**	The wolf and the crocodile were reading a book.
365	**F**	**114**	The teacher and the student were reading a book.
366	**F**	**172**	The squirrel and the pigeon were flying a kite.
367	**T**	**17**	Sandra brought the... the fly to an octopus.
368	**F**	**224**	The penguin and the eagle were camping.
369	**F**	**130**	Charles was hiking.
370	**F**	**126**	The hippo and the mole were hiking.
371	**T**	**18**	The guard brought the... the butterfly to a snail.
372	**F**	**112**	The sheep and the turkey were reading a book.
373	**F**	**219**	The zebra and the crocodile were juggling.
374	**F**	**94**	The eagle and the pelican were watching tv.
375	**T**	**13**	Anthony lent the... the ant to a panda.
376	**F**	**97**	The hamster and the rat were watching tv.
377	**F**	**128**	Lisa and Daniel were hiking.
378	**F**	**102**	The owl and the dog were doing shopping.
379	**T**	**14**	The carpenter lent the... the mole to a shark.
380	**F**	**225**	The giraffe and the moth were camping.
381	**F**	**94**	The eagle and the pelican were watching tv.
382	**F**	**210**	The wolf and the frog were jumping up and down.
383	**T**	**15**	George showed the... the duck to an eagle.
384	**F**	**72**	Paul and Raul were cooking.
385	**F**	**142**	The horse and the bull were cleaning.
386	**F**	**184**	The panther and the monkey were skateboarding.
387	**T**	**19**	Helen awarded the... the snake to a rooster.

388	**F**	**153**	The fairy was playing the guitar.
389	**F**	**120**	Isabella was reading a book.
390	**F**	**144**	The seal and the swan were cleaning.
391	**T**	**16**	The cowgirl showed the... the fish to an elephant.
392	**F**	**227**	The hyena and the toucan were camping.
393	**F**	**124**	The pelican and the crab were hiking.
394	**F**	**61**	The ant and cow were having dinner.
395	**T**	**12**	The actress faxed the... the hamster to a swan.
396	**F**	**118**	The bear and the butterfly were reading a book.
397	**F**	**160**	Sandra was playing the guitar.
398	**F**	**133**	The rhino and the hamster were painting.
399	**T**	**11**	Charles faxed the... the owl to a ray.
400	**F**	**85**	The bandit and the barista were playing tennis.